

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 5, 2006, 20:53:35 ; Search time 46 Seconds  
(without alignments)  
666.797 Million cell updates/sec

Title: US-09-872-505A-6  
Perfect score: 1892  
Sequence: 1 MITMLWHAMPELNTARLMA.....PLAQRERDEDDWDEEDW 371

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/iaa/5 COMB.pep.\*  
2: /cgn2\_6/prodata/1/iaa/6 COMB.pep.\*  
3: /cgn2\_6/prodata/1/iaa/H COMB.pep.\*  
4: /cgn2\_6/prodata/1/iaa/PCUS COMB.pep.\*  
5: /cgn2\_6/prodata/1/iaa/RE COMB.pep.\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1892	100.0	371	2	US-09-050-739-92
2	1873	99.0	368	2	US-08-818-112-114
3	1873	99.0	368	2	US-08-818-111-109
4	1873	99.0	368	2	US-09-056-556-114
5	1873	99.0	368	2	US-09-072-596-109
6	1873	99.0	368	2	US-09-072-967-114
7	1873	99.0	368	2	US-10-193-002-109
8	1873	99.0	368	2	US-10-084-843-114
9	366	19.3	400	2	US-09-073-009-126
10	366	19.3	400	2	US-09-073-010-126
11	344.5	18.2	394	2	US-09-712-363-205
12	330.5	17.5	729	2	US-09-223-040-2
13	330.5	17.5	729	2	US-09-287-849-2
14	328.5	17.4	596	2	US-09-287-849-26
15	326	17.2	423	2	US-09-073-009-142
16	326	17.2	423	2	US-09-073-010-142
17	326	17.2	710	2	US-09-287-849-16
18	326	17.2	856	2	US-09-287-849-12
19	318	16.8	600	2	US-08-818-112-107
20	317	16.8	391	2	US-08-818-111-102
21	317	16.8	391	2	US-08-818-111-102
22	317	16.8	391	2	US-09-056-556-107
23	317	16.8	391	2	US-09-072-596-102
24	317	16.8	391	2	US-09-072-967-107
25	317	16.8	391	2	US-10-193-002-102
26	317	16.8	391	2	US-10-084-843-107
27	312	16.5	396	2	US-08-818-112-111

28	312	16.5	396	2	US-08-818-111-106	Sequence 106, App
29	312	16.5	396	2	US-09-056-556-111	Sequence 111, App
30	312	16.5	396	2	US-09-072-596-106	Sequence 106, App
31	312	16.5	396	2	US-09-072-967-111	Sequence 111, App
32	312	16.5	396	2	US-10-193-002-106	Sequence 106, App
33	312	16.5	396	2	US-10-084-843-111	Sequence 111, App
34	311.5	16.5	1271	1	US-08-095-734-2	Sequence 2, Appli
35	311.5	16.5	1271	1	US-08-444-623-2	Sequence 2, Appli
36	311.5	16.5	1271	2	US-08-471-869-2	Sequence 2, Appli
37	311.5	16.5	1271	2	US-09-342-563-2	Sequence 2, Appli
38	311.5	16.5	1271	4	PCT-US94-08267-2	Sequence 109, App
39	294.5	15.6	359	2	US-08-818-112-109	Sequence 104, App
40	294.5	15.6	359	2	US-08-818-111-104	Sequence 109, App
41	294.5	15.6	359	2	US-09-056-556-109	Sequence 104, App
42	294.5	15.6	359	2	US-09-072-596-104	Sequence 109, App
43	294.5	15.6	359	2	US-09-072-967-109	Sequence 104, App
44	294.5	15.6	359	2	US-10-193-002-104	Sequence 109, App
45	294.5	15.6	359	2	US-10-084-843-109	Sequence 109, App

RESULT 1  
US-09-050-739-92  
; Sequence 92, Application US/09050739  
; Patent No. 6641814  
; GENERAL INFORMATION:  
; APPLICANT: ANDERSEN, Peter  
; APPLICANT: NIELSEN, Rikke  
; APPLICANT: OETTINGER, Thomas  
; APPLICANT: RASMUSSEN, Peter Birk  
; APPLICANT: ROSENKRANDS, Ida  
; APPLICANT: WELDINGH, Karin  
; APPLICANT: FLORIO, Walter  
; TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS  
; TITLE OF INVENTION: DERIVED FROM M. TUBERCULOSIS  
; FILE REFERENCE: 670001-2002.1  
; CURRENT APPLICATION NUMBER: US/09/050,739  
; CURRENT FILING DATE: 1998-03-30  
; EARLIER APPLICATION NUMBER: 0376/97  
; EARLIER FILING DATE: 1997-04-02  
; EARLIER APPLICATION NUMBER: 1277/97  
; EARLIER FILING DATE: 1997-11-10  
; EARLIER APPLICATION NUMBER: 60/044,624  
; EARLIER FILING DATE: 1997-04-18  
; EARLIER APPLICATION NUMBER: 60/070,488  
; EARLIER FILING DATE: 1998-01-05  
; NUMBER OF SEQ ID NOS: 173  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 92  
; LENGTH: 371  
; TYPE: PRT  
; ORGANISM: Mycobacterium tuberculosis  
US-09-050-739-92

Query Match 100.0%; Score 1892; DB 2; Length 371;  
Best Local Similarity 100.0%; Pred. No. 1.4e-154;  
Matches 371; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MITMLWHAMPELNTARLMAGAGPAPMLAAAGWQTLAALDAQAVELTARLNSLGEAWT	60
Db	1	MITMLWHAMPELNTARLMAGAGPAPMLAAAGWQTLAALDAQAVELTARLNSLGEAWT	60
QY	61	GGGSDKALAAATPMVWLQASTQAKTRAMQATAQAAATQAAATTPSLPEIAANHITQAA	120
Db	61	GGGSDKALAAATPMVWLQASTQAKTRAMQATAQAAATQAAATTPSLPEIAANHITQAA	120
QY	121	VLTAATFFGINTIPIALTENDYFIRWNNQAALAMEVYQAEAVNTLFEKLEPMASILDPG	180
Db	121	VLTAATFFGINTIPIALTENDYFIRWNNQAALAMEVYQAEAVNTLFEKLEPMASILDPG	180
QY	181	ASQSTTNPIFGMPSPGSSTPVGQLPPAATQTLGLGEMSGPMQOLTQPLQOQVTSLSFQVG	240

ALIGNMENTS

Db 181 ASOSTNPIFGWSPGSGSTPVGQLPPAATQTGLGEMSGPMQQLTQPLQQVTSLSFQVG 240  
Qy 241 GTGGGNPADEEAAQOMGLLGTSPLSNHPLAGSGSPSAGAGLLRAESLPGAGGSLTRTPLMS 300  
Db 241 GTGGGNPADEEAAQOMGLLGTSPLSNHPLAGSGSPSAGAGLLRAESLPGAGGSLTRTPLMS 300  
Qy 301 QLIKXPVAPSVMPAAAGSSATGGAAPVCGAGAMGQAGSGSTRPGLVAPAPLAQERED 360  
Db 301 QLIKXPVAPSVMPAAAGSSATGGAAPVCGAGAMGQAGSGSTRPGLVAPAPLAQERED 360  
Qy 361 DEDDWEDEDDW 371  
Db 361 DEDDWEDEDDW 371  
RESULT 2  
US-08-818-112-114  
; Sequence 114, Application US/08818112  
; Patent No. 6290969  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Campos-Neto, Antonio  
; APPLICANT: Houghton, Raymond  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Twardzik, Daniel R.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF TUBERCULOSIS  
; NUMBER OF SEQUENCES: 153  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/818,112  
; FILING DATE: 13-MAR-1997  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maki, David J.  
; REGISTRATION NUMBER: 31,392  
; REFERENCE/DOCKET NUMBER: 210121.411C6  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 114:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 368 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-818-112-114

Query Match 99.0%; Score 1873; DB 2; Length 368;  
Best Local Similarity 99.7%; Pred. No. 6e-153;  
Matches 367; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 MLWHAMPELNTARLMAGAGPAPMLAAAGQWTLSSAALDAQAVELTARLNSLGEAWTGGG 63  
Db 1 MLWHAMPEXNTARLMAGAGPAPMLAAAGQWTLSSAALDAQAVELTARLNSLGEAWTGGG 60  
Qy 64 SDKALAAATPMVWMLQTAQTAKTRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 123  
Db 61 SDKALAAATPMVWMLQTAQTAKTRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 120

Qy 124 ATNPFGINTIPIALTEMDFIRMNQOALAMEVYQAEAVNTLFEKLEPMASILDPGASQ 183  
Db 121 ATNPFGINTIPIALTEMDFIRMNQOALAMEVYQAEAVNTLFEKLEPMASILDPGASQ 180  
Qy 184 STTNPIFCMPSPGSGSTPVGQLPPAATQTGLGEMSGPMQQLTQPLQQVTSLSFQVG 243  
Db 181 STTNPIFCMPSPGSGSTPVGQLPPAATQTGLGEMSGPMQQLTQPLQQVTSLSFQVG 240  
Qy 244 GGNPADBEAAQOMGLLGTSPLSNHPLAGSGSPSAGAGLLRAESLPGAGGSLTRTPLMS 303  
Db 241 GGNPADBEAAQOMGLLGTSPLSNHPLAGSGSPSAGAGLLRAESLPGAGGSLTRTPLMS 300  
Qy 304 EKPVPAPSVMPAAAGSSATGGAAPVCGAGAMGQAGSGSTRPGLVAPAPLAQERED 363  
Db 301 EKPVPAPSVMPAAAGSSATGGAAPVCGAGAMGQAGSGSTRPGLVAPAPLAQERED 360  
Qy 364 DWEDEDDW 371  
Db 361 DWEDEDDW 368  
RESULT 3  
US-08-818-111-109  
; Sequence 109, Application US/08818111  
; Patent No. 6338852  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Campos-Neto, Antonio  
; APPLICANT: Houghton, Raymond  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Twardzik, Daniel R.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF  
; NUMBER OF SEQUENCES: 148  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/818,111  
; FILING DATE: 13-MAR-1997  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maki, David J.  
; REGISTRATION NUMBER: 31,392  
; REFERENCE/DOCKET NUMBER: 210121.417C6  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 109:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 368 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-818-111-109

Query Match 99.0%; Score 1873; DB 2; Length 368;  
Best Local Similarity 99.7%; Pred. No. 6e-153;  
Matches 367; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 MLWHAMPELNTARLMAGAGPAPMLAAAGQWTLSSAALDAQAVELTARLNSLGEAWTGGG 63  
Db 1 MLWHAMPEXNTARLMAGAGPAPMLAAAGQWTLSSAALDAQAVELTARLNSLGEAWTGGG 60

```

Db 1 MLWAMPPEXNTARLMAGAGPAPMLAAAGWQTLSSAALDAQAVELTARLNSLGEAWTGGG 60
Qy 64 SDKALAAATPMVWLQTAQAKTRAMQATQAQAAAYTOAMATTPLSLPEIAANHITQAVLT 123
Db 61 SDKALAAATPMVWLQTAQAKTRAMQATQAQAAAYTOAMATTPLSLPEIAANHITQAVLT 120
Qy 124 ATNFFGINTPIALTETMDYFIRMNQALAMEVYQAEAVNTLPEKLEPMASILDPGASQ 183
Db 121 ATNFFGINTPIALTETMDYFIRMNQALAMEVYQAEAVNTLPEKLEPMASILDPGASQ 180
Qy 184 STTNPIFGMPSPGSGSTPVGQLPPAATQTLGQGENSGPMQQLTQPLQVTSLSFQVGGTG 243
Db 181 STTNPIFGMPSPGSGSTPVGQLPPAATQTLGQGENSGPMQQLTQPLQVTSLSFQVGGTG 240
Qy 244 GGNPADEEAAQWGLLGTSPLSNHPPLAGGSGPSAGAGLLRAESLPGAGGSLTRTPLMSQLI 303
Db 241 GGNPADEEAAQWGLLGTSPLSNHPPLAGGSGPSAGAGLLRAESLPGAGGSLTRTPLMSQLI 300
Qy 304 EKPVPASVMPAAAAGSSATGGAAPVGAGAMGQAGSGGSTRPGLVAPAPLAQEREDDED 363
Db 301 EKPVPASVMPAAAAGSSATGGAAPVGAGAMGQAGSGGSTRPGLVAPAPLAQEREDDED 360
Qy 364 DWDEEDDW 371
Db 361 DWDEEDDW 368

RESULT 4
US-09-056-556-114
; Sequence 114, Application US/09056556
; Patent No. 6350456
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND
; NUMBER OF SEQUENCES: 241
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 07-APR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.457
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 114:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 368 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-09-056-556-114
Query Match 99.0%; Score 1873; DB 2; Length 368;
Best Local Similarity 99.7%; Pred. No. 6e-153;
Matches 367; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 MLWAMPPEXNTARLMAGAGPAPMLAAAGWQTLSSAALDAQAVELTARLNSLGEAWTGGG 63

```

```

Db 1 MLWAMPPEXNTARLMAGAGPAPMLAAAGWQTLSSAALDAQAVELTARLNSLGEAWTGGG 60
Qy 64 SDKALAAATPMVWLQTAQAKTRAMQATQAQAAAYTOAMATTPLSLPEIAANHITQAVLT 123
Db 61 SDKALAAATPMVWLQTAQAKTRAMQATQAQAAAYTOAMATTPLSLPEIAANHITQAVLT 120
Qy 124 ATNFFGINTPIALTETMDYFIRMNQALAMEVYQAEAVNTLPEKLEPMASILDPGASQ 183
Db 121 ATNFFGINTPIALTETMDYFIRMNQALAMEVYQAEAVNTLPEKLEPMASILDPGASQ 180
Qy 184 STTNPIFGMPSPGSGSTPVGQLPPAATQTLGQGENSGPMQQLTQPLQVTSLSFQVGGTG 243
Db 181 STTNPIFGMPSPGSGSTPVGQLPPAATQTLGQGENSGPMQQLTQPLQVTSLSFQVGGTG 240
Qy 244 GGNPADEEAAQWGLLGTSPLSNHPPLAGGSGPSAGAGLLRAESLPGAGGSLTRTPLMSQLI 303
Db 241 GGNPADEEAAQWGLLGTSPLSNHPPLAGGSGPSAGAGLLRAESLPGAGGSLTRTPLMSQLI 300
Qy 304 EKPVPASVMPAAAAGSSATGGAAPVGAGAMGQAGSGGSTRPGLVAPAPLAQEREDDED 363
Db 301 EKPVPASVMPAAAAGSSATGGAAPVGAGAMGQAGSGGSTRPGLVAPAPLAQEREDDED 360
Qy 364 DWDEEDDW 371
Db 361 DWDEEDDW 368

RESULT 5
US-09-072-596-109
; Sequence 109, Application US/09072596
; Patent No. 6458366
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonia
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
; NUMBER OF SEQUENCES: 350
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 05-MAY-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.417C9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 109:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 368 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
TUBERCULOSIS

```

US-09-072-596-109

Query Match 99.0%; Score 1873; DB 2; Length 368;  
Best Local Similarity 99.7%; Pred. No. 6e-153;  
Matches 367; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 4 MLWHAMPPELNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWTGGG 63  
DB 1 MLWHAMPPEXNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWTGGG 60  
QY 64 SDKALAAATPMVWLQASTQAKTRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 123  
DB 61 SDKALAAATPMVWLQASTQAKTRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 120  
QY 124 ATNFFGINTPIALTMDYFIRMNQALAMEVYQAEAVNTLFEKLEPMASILDPGASQ 183  
DB 121 ATNFFGINTPIALTMDYFIRMNQALAMEVYQAEAVNTLFEKLEPMASILDPGASQ 180  
QY 184 STTNPIFGMPSPGSGSTPVGQLPPAATOTLQGLGEMSGPMQQLTOPLOQVTSLSFQVGGTG 243  
DB 181 STTNPIFGMPSPGSGSTPVGQLPPAATOTLQGLGEMSGPMQQLTOPLOQVTSLSFQVGGTG 240  
QY 244 GGNPADEEAAQMLLGTSPLSNHLPLAGSGSPSAGAGLLRAESLPGAGGSLTRTPLMSQLI 303  
DB 241 GGNPADEEAAQMLLGTSPLSNHLPLAGSGSPSAGAGLLRAESLPGAGGSLTRTPLMSQLI 300  
QY 304 EKPVPAPSVMPAAAAGSSATGGAAPVAGAGMGQAGSGSTRPGLVAPAPLAQEREDEDED 363  
DB 301 EKPVPAPSVMPAAAAGSSATGGAAPVAGAGMGQAGSGSTRPGLVAPAPLAQEREDEDED 360  
QY 364 DWDEEDDW 371  
DB 361 DWDEEDDW 368

RESULT 6

US-09-072-967-114  
; Sequence 114, Application US/09072967  
; Patent No. 6592877  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Campos-Neto, Antonio  
; APPLICANT: Houghton, Raymond  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Twardzik, Daniel R.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Hendrickson, Ronald C.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY  
; NUMBER OF SEQUENCES: 355  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/072,967  
; FILING DATE: 05-MAY-1998  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maki, David J.  
; REGISTRATION NUMBER: 31,392  
; REFERENCE/DOCKET NUMBER: 210121.411C9  
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 114:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 368 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-09-072-967-114

Query Match 99.0%; Score 1873; DB 2; Length 368;  
Best Local Similarity 99.7%; Pred. No. 6e-153;  
Matches 367; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 4 MLWHAMPPELNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWTGGG 63  
DB 1 MLWHAMPPEXNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWTGGG 60  
QY 64 SDKALAAATPMVWLQASTQAKTRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 123  
DB 61 SDKALAAATPMVWLQASTQAKTRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 120  
QY 124 ATNFFGINTPIALTMDYFIRMNQALAMEVYQAEAVNTLFEKLEPMASILDPGASQ 183  
DB 121 ATNFFGINTPIALTMDYFIRMNQALAMEVYQAEAVNTLFEKLEPMASILDPGASQ 180  
QY 184 STTNPIFGMPSPGSGSTPVGQLPPAATOTLQGLGEMSGPMQQLTOPLOQVTSLSFQVGGTG 243  
DB 181 STTNPIFGMPSPGSGSTPVGQLPPAATOTLQGLGEMSGPMQQLTOPLOQVTSLSFQVGGTG 240  
QY 244 GGNPADEEAAQMLLGTSPLSNHLPLAGSGSPSAGAGLLRAESLPGAGGSLTRTPLMSQLI 303  
DB 241 GGNPADEEAAQMLLGTSPLSNHLPLAGSGSPSAGAGLLRAESLPGAGGSLTRTPLMSQLI 300  
QY 304 EKPVPAPSVMPAAAAGSSATGGAAPVAGAGMGQAGSGSTRPGLVAPAPLAQEREDEDED 363  
DB 301 EKPVPAPSVMPAAAAGSSATGGAAPVAGAGMGQAGSGSTRPGLVAPAPLAQEREDEDED 360  
QY 364 DWDEEDDW 371  
DB 361 DWDEEDDW 368

RESULT 7

US-10-193-002-109  
; Sequence 109, Application US/10193002  
; Patent No. 6949246  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Campos-Neto, Antonio  
; APPLICANT: Houghton, Raymond  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Twardzik, Daniel R.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Hendrickson, Ronald C.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF  
; TUBERCULOSIS  
; NUMBER OF SEQUENCES: 350  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/193,002  
 FILING DATE: 10-Jul-2002  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/09/072,596  
 FILING DATE: 05-MAY-1998  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Maki, David J.  
 REGISTRATION NUMBER: 31,392  
 REFERENCE/DOCKET NUMBER: 210121.417C9  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (206) 622-4900  
 TELEFAX: (206) 682-6031  
 INFORMATION FOR SEQ ID NO: 109:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 368 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 SEQUENCE DESCRIPTION: SEQ ID NO: 109:  
 US-10-193-002-109

Query Match 99.0%; Score 1873; DB 2; Length 368;  
 Best Local Similarity 99.7%; Pred. No. 6e-153;  
 Matches 367; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 MLWHAMPPELNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWTGGG 63  
 DB 1 MLWHAMPPEXNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWTGGG 60  
 QY 64 SKALAAATPMVWLQTAQAKRAMQATAQAAYTQAMATTPSLPEIAANHITQAVLT 123  
 DB 61 SDKALAAATPMVWLQTAQAKRAMQATAQAAYTQAMATTPSLPEIAANHITQAVLT 120  
 QY 124 ATNFFGINTPIALTETMDYFIRMNQALAMEVYQAEVAVNTLFEKLEPMASILDPGASQ 183  
 DB 121 ATNFFGINTPIALTETMDYFIRMNQALAMEVYQAEVAVNTLFEKLEPMASILDPGASQ 180  
 QY 184 STTNPIFGMPSGSSSTPVQQLPPAATQTLLGQEMSGPMQQLTPQLQVTSLSFSQVGGTG 243  
 DB 181 STTNPIFGMPSGSSSTPVQQLPPAATQTLLGQEMSGPMQQLTPQLQVTSLSFSQVGGTG 240  
 QY 244 GGNPADEEAQMGLLGTSPLSNHPLAGGSPSAGAGLLRAESLPGAGGSLTRTPPLMSQLI 303  
 DB 241 GGNPADEEAQMGLLGTSPLSNHPLAGGSPSAGAGLLRAESLPGAGGSLTRTPPLMSQLI 300  
 QY 304 EKPVPASVMPAAAGSSATGGAAPVGAGAMGQAGQSGGSTRPGLVAPAPLAQREDEDED 363  
 DB 301 EKPVPASVMPAAAGSSATGGAAPVGAGAMGQAGQSGGSTRPGLVAPAPLAQREDEDED 360  
 QY 364 DWDEEDDW 371  
 DB 361 DWDEEDDW 368

RESULT 8  
 US-10-084-843-114  
 ; Sequence 114, Application US/10084843  
 ; Patent No. 6962710  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Reed, Steven G.  
 ; Skeiky, Yasir A.W.  
 ; Dillon, Davin C.  
 ; Campos-Neto, Antonio  
 ; Houghton, Raymond  
 ; Vedvick, Thomas S.  
 ; Twardzik, Daniel R.  
 ; Lodes, Michael J.  
 ; Hendrickson, Ronald C.  
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY  
 ; NUMBER OF SEQUENCES: 355  
 ; CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP  
 STREET: 6300 Columbia Center, 701 Fifth Avenue  
 CITY: Seattle  
 STATE: Washington  
 COUNTRY: USA  
 ZIP: 98104-7092  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/084,843  
 FILING DATE: 25-Feb-2002  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/09/072,967  
 FILING DATE: 05-MAY-1998  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Maki, David J.  
 REGISTRATION NUMBER: 31,392  
 REFERENCE/DOCKET NUMBER: 210121.411C9  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (206) 622-4900  
 TELEFAX: (206) 682-6031  
 INFORMATION FOR SEQ ID NO: 114:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 368 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 SEQUENCE DESCRIPTION: SEQ ID NO: 114:  
 US-10-084-843-114

Query Match 99.0%; Score 1873; DB 2; Length 368;  
 Best Local Similarity 99.7%; Pred. No. 6e-153;  
 Matches 367; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 MLWHAMPPELNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWTGGG 63  
 DB 1 MLWHAMPPEXNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWTGGG 60  
 QY 64 SKALAAATPMVWLQTAQAKRAMQATAQAAYTQAMATTPSLPEIAANHITQAVLT 123  
 DB 61 SKALAAATPMVWLQTAQAKRAMQATAQAAYTQAMATTPSLPEIAANHITQAVLT 120  
 QY 124 ATNFFGINTPIALTETMDYFIRMNQALAMEVYQAEVAVNTLFEKLEPMASILDPGASQ 183  
 DB 121 ATNFFGINTPIALTETMDYFIRMNQALAMEVYQAEVAVNTLFEKLEPMASILDPGASQ 180  
 QY 184 STTNPIFGMPSGSSSTPVQQLPPAATQTLLGQEMSGPMQQLTPQLQVTSLSFSQVGGTG 243  
 DB 181 STTNPIFGMPSGSSSTPVQQLPPAATQTLLGQEMSGPMQQLTPQLQVTSLSFSQVGGTG 240  
 QY 244 GGNPADEEAQMGLLGTSPLSNHPLAGGSPSAGAGLLRAESLPGAGGSLTRTPPLMSQLI 303  
 DB 241 GGNPADEEAQMGLLGTSPLSNHPLAGGSPSAGAGLLRAESLPGAGGSLTRTPPLMSQLI 300  
 QY 304 EKPVPASVMPAAAGSSATGGAAPVGAGAMGQAGQSGGSTRPGLVAPAPLAQREDEDED 363  
 DB 301 EKPVPASVMPAAAGSSATGGAAPVGAGAMGQAGQSGGSTRPGLVAPAPLAQREDEDED 360  
 QY 364 DWDEEDDW 371  
 DB 361 DWDEEDDW 368

RESULT 9  
 US-09-073-009-126  
 ; Sequence 126, Application US/09073009  
 ; Patent No. 6555653  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Alderson, Mark





```
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:tri-fusion
US-09-287-849-2

Query Match      17.5%; Score 330.5; DB 2; Length 729;
Best Local Similarity 28.9%; Pred. No. 6.8e-20;
Matches 120; Conservative 45; Mismatches 173; Indels 77; Gaps 13;

QY      8  AMPPELNTARLMAGAGPAPMLAAAAGWQTL-----SAALDAQAVELTARLNSLGEAWTGG 62
Db      147  ALPPEINSARMYAGPGSASLVAAAQMWDSVASDLFSAASAFQSVVMGLTVGS-----WIGS 202

QY      63  GSDKALAAATPMVWLQTAQKTRAMQATAQAAAAYTOAMATTTPSLPEIAANHITQAVL 122
Db      203  SAGLMVAASPYVAMSVTAGQBELTAAQVRVAAAAYETAYGLTVPPPVIAENRAELMIL 262

QY      123  TATNFFGINTIPALTENDYFIRMNQALAMEVYQAEAVNT----- 165
Db      263  IATNLGQNTPAIAVNEAEG-EMWQDAAAMFGYAAATATATATATLLPPEEAPEMTSAGG 321

QY      166  LFEKLEPMASILDPCASQSTTNPI-----FCMPSPGSSTPVGQL-----PPAATQTL 212
Db      322  LLEQAAAVEEASDTAAANQLMNNVPQALQQLAQPTQG-TTPSSKLGGLWKTVPSPHRSPIS 380

QY      213  GOLGEMSGPMQQLTQPLQOVTSLSFQVGG--TGGGNPADEEAAQWGLLGTSPLSNHPLAG 270
Db      381  NNVSWMNNHMSMTNSGVSMNTLSMLKGFAPAAARQAVQTAQNGVVRAMSSLGSSLSGS 440

QY      271  GSGPSAGAGLLRAESLPGAGGSLTRTPLSQLIEKPVP-----SVMPAAAAAGSSAT 322
Db      441  GLGGVVAANLGRAASV-----GSLs-VPQAWAAAANQAVTPAARALPLTSLTSAERGPQM 495

QY      323  GGAAPVAGAMGQQAQSGG-----STRPGLVAPAPLAQREED 360
Db      496  LGGLPVQG----MGARAGGGLSGVLRVPPRPYVMPHSPAAGDIAPPALSDQRFAD 546
```

```
RESULT 14
US-09-287-849-26
; Sequence 26, Application US/09287849
; Patent No. 6627198
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Alderson, Mark
; APPLICANT: Campos-Neto, Antonio
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
; FILE REFERENCE: 014058-009020US
; CURRENT APPLICATION NUMBER: US/09/287,849
; CURRENT FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: US 08/818,112
; PRIOR FILING DATE: 1997-03-13
; PRIOR APPLICATION NUMBER: US 08/942,578
; PRIOR FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: US 09/025,197
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 09/056,556
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: US 09/223,040
```

```
; PRIOR FILING DATE: 1998-12-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 596
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:bi-fusion
US-09-287-849-26
```

```
Query Match      17.4%; Score 328.5; DB 2; Length 596;
Best Local Similarity 28.9%; Pred. No. 7.7e-20;
Matches 120; Conservative 45; Mismatches 173; Indels 77; Gaps 13;

QY      8  AMPPELNTARLMAGAGPAPMLAAAAGWQTL-----SAALDAQAVELTARLNSLGEAWTGG 62
Db      14  ALPPEINSARMYAGPGSASLVAAAQMWDSVASDLFSAASAFQSVVMGLTVGS-----WIGS 69

QY      63  GSDKALAAATPMVWLQTAQKTRAMQATAQAAAAYTOAMATTTPSLPEIAANHITQAVL 122
Db      70  SAGLMVAASPYVAMSVTAGQBELTAAQVRVAAAAYETAYGLTVPPPVIAENRAELMIL 129

QY      123  TATNFFGINTIPALTENDYFIRMNQALAMEVYQAEAVNT----- 165
Db      130  IATNLGQNTPAIAVNEAEG-EMWQDAAAMFGYAAATATATATATLLPPEEAPEMTSAGG 188

QY      166  LFEKLEPMASILDPCASQSTTNPI-----FCMPSPGSSTPVGQL-----PPAATQTL 212
Db      189  LLEQAAAVEEASDTAAANQLMNNVPQALQQLAQPTQG-TTPSSKLGGLWKTVPSPHRSPIS 247

QY      213  GOLGEMSGPMQQLTQPLQOVTSLSFQVGG--TGGGNPADEEAAQWGLLGTSPLSNHPLAG 270
Db      248  NNVSWMNNHMSMTNSGVSMNTLSMLKGFAPAAARQAVQTAQNGVVRAMSSLGSSLSGS 307

QY      271  GSGPSAGAGLLRAESLPGAGGSLTRTPLSQLIEKPVP-----SVMPAAAAAGSSAT 322
Db      308  GLGGVVAANLGRAASV-----GSLs-VPQAWAAAANQAVTPAARALPLTSLTSAERGPQM 362

QY      323  GGAAPVAGAMGQQAQSGG-----STRPGLVAPAPLAQREED 360
Db      363  LGGLPVQG----MGARAGGGLSGVLRVPPRPYVMPHSPAAGDIAPPALSDQRFAD 413
```

```
RESULT 15
US-09-073-009-142
; Sequence 142, Application US/09073009
; Patent No. 6555653
; GENERAL INFORMATION:
; APPLICANT: Alderson, Mark
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Campos-Neto, Antonio
; TITLE OF INVENTION: COMPOUNDS AND DIAGNOSIS OF
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY
; STREET: 6300 Columbia Center, 701 Fifth Ave.
; CITY: Seattle
; STATE: Washington
; COUNTRY: US
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,009
; FILING DATE: 05-MAY-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
```

TUBERCULOSIS AND MR



REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.441C1  
TELEPHONE: 206-622-4900  
TELEFAX: 206-682-6031  
INFORMATION FOR SEQ ID NO: 142:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 423 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-073-009-142

Query Match 17.2%; Score 326; DB 2; Length 423;  
Best Local Similarity 30.4%; Pred. NO. 7.9e-20;  
Matches 121; Conservative 45; Mismatches 164; Indels 68; Gaps 17;

QY	9	MPELNTARLMAGAGPAPMLAAAGWQTLSSAALDAQAVELTARLNSL-GEAWTGGSDKA	67
Db	6	LPPEVNSRMVSGPGPEKSLAAAANDGVAAELTSAAYSGVSVTLIVEPMWGFAAAM	65
QY	68	LAATPMVVLQSTASTOAKTRAMQATAQAAYTQAMATT--PSLPEIAANHITOAVLTAT	125
Db	66	AAATPYVGLAATAALAKETATQAAAAEAFTAFAMTVPESL--VAANRSLMSLVAA	123
QY	126	NFFGINTIPIALTEMDFIRMNQALAMEVYOETAVENTLFEKLEPM---ASILDPGAS	182
Db	124	NILQNSAAIATQAEY-AEMWAQDAAVMYSYEGASAAASALPPFTPPVQGTGPAGPAA	182
QY	183	QSTTNPIFGMPSPGSGSTPVGQLPP-----AATQTLGOLGEMS--GPMQQLTQ	227
Db	183	AAATQAAGAGAVADAQATLAQLPPGILSDILSALAANDPLTSGLLGIASTINPQVGS	242
QY	228	PL-----QQVTSLSFQVGGTGGGNPADEBAQ---MGLLGT-----PLSNHPLA---	269
Db	243	PVITPTIGELDVIALYIASIATGSIATITWTRPWHIGLYGNAGGLGPTQGHPLSSAT	302
QY	270	-----GGSGP-SAGAG---LLRAESLPGAGGSLTRTPLMSQLIEKVPVPSVMPAA	315
Db	303	DEPEPHWGPFPGGAAPVSGVGHAAVLGALSVPHSW--TTAAPEIQLAVQ--ATPTFSSA	358
QY	316	AGSSATGCAAPVG-----AGAMGQGAQSGGSTRPG	346
Db	359	GADPTALNG-MPAGLLSGMALASLAARGTTGGGTRSG	395

Search completed: January 5, 2006, 21:19:28  
Job time : 47 secs

This Page Blank (usp10)

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 5, 2006, 16:38:03 ; Search time 138 Seconds  
(without alignments)  
1181.228 Million cell updates/sec

Title: US-09-872-505A-6  
Perfect score: 1892  
Sequence: 1 MITMLHAMPPELNTARLMA.....PLAQREREDDDWDEEDDW 371

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_21.\*  
1: geneseqp1980s.\*  
2: geneseqp1990s.\*  
3: geneseqp2000s.\*  
4: geneseqp2001s.\*  
5: geneseqp2002s.\*  
6: geneseqp2003as.\*  
7: geneseqp2003bs.\*  
8: geneseqp2004s.\*  
9: geneseqp2005s.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1892	100.0	371	AAW72929	Mycobacte
2	1892	100.0	371	AAV21946	Aay21946 Amino aci
3	1878	99.3	368	AAV19844	AAV19844 Mycobacte
4	1878	99.3	368	ADG15733	Adg15733 M tubercu
5	1873	99.0	368	AAW32384	AAW32384 Mycobacte
6	1873	99.0	368	AAW32452	AAW32452 Mycobacte
7	1873	99.0	368	AAW43338	AAW43338 Mycobacte
8	1873	99.0	368	AAW81705	AAW81705 M. tuberc
9	1873	99.0	368	AAV38992	AAV38992 M. tuberc
10	1873	99.0	368	AAV39135	AAV39135 M. tuberc
11	1873	99.0	368	AAV05362	AAV05362 M. tuberc
12	502	26.5	100	ADG15750	Adg15750 Rv3873 se
13	436.5	23.1	445	ABU34364	Abu34364 Protein e
14	430.5	22.8	434	AAW66451	AAW66451 Protein e
15	430.5	22.8	434	ABU36905	ABU36905 Protein e
16	366	19.3	400	AAW73654	AAW73654 M. tuberc
17	366	19.3	400	AAW73764	AAW73764 M. tuberc
18	366	19.3	400	AAU08225	AAU08225 Mycobacte
19	363.5	19.2	405	ABU34103	ABU34103 Protein e
20	362	19.1	408	ABM15864	ABM15864 Mycobacte
21	356	18.8	572	ABU36014	ABU36014 Protein e
22	355.5	18.8	467	ABU34048	ABU34048 Protein e
23	344.5	18.2	394	AAW81154	AAW81154 Mycobacte
24	343	18.1	416	ADZ36102	ADZ36102 Mycobacte

25	333.5	17.6	729	7	ADA26373	Ada26373 Mycobacte
26	330.5	17.5	729	5	AAE29708	Aae29708 Mycobacte
27	330.5	17.5	729	5	AAE17572	Aae17572 Mycobacte
28	330.5	17.5	729	7	ADF69753	Adf69753 Fusion pr
29	330.5	17.5	1010	7	ADA26356	Ada26356 Mycobacte
30	330	17.4	1154	7	ADA26368	Ada26368 Mycobacte
31	328.5	17.4	596	2	AAV32070	Aav32070 Mycobacte
32	328.5	17.4	596	5	AAE29710	Aae29710 Mycobacte
33	328.5	17.4	596	5	AAE17574	Aae17574 Mycobacte
34	328.5	17.4	596	7	ADF69777	Adf69777 Fusion pr
35	328.5	17.4	599	5	AAU74599	Aau74599 Antigenic
36	328.5	17.4	729	4	AAO22142	Aao22142 Rai2-H9-3
37	328.5	17.4	729	5	AAE29709	Aae29709 Mycobacte
38	328.5	17.4	729	5	AAE17573	Aae17573 Mycobacte
39	328.5	17.4	729	7	ADA26374	Ada26374 Mycobacte
40	328.5	17.4	813	7	ADA26367	Ada26367 Mycobacte
41	328.5	17.4	825	7	ADA26366	Ada26366 Mycobacte
42	328.5	17.4	875	7	ADA26365	Ada26365 Mycobacte
43	328.5	17.4	930	5	AAE29731	Aae29731 Mycobacte
44	328.5	17.4	930	7	ADA26364	Ada26364 Mycobacte
45	328.5	17.4	1016	7	ADA26370	Ada26370 M. bovis

## ALIGNMENTS

RESULT 1	
AAW72929	
ID	AAW72929 standard; protein; 371 AA.
XX	
AC	AAW72929;
XX	
DT	21-JAN-1999 (first entry)
XX	
DE	Mycobacterium tuberculosis antigen RD1-ORF5.
XX	
KW	Mycobacterium tuberculosis; antigen; vaccine; immunological; immunogen;
KW	infection.
XX	
OS	Mycobacterium tuberculosis.
XX	
PN	W09844119-A1.
XX	
PD	08-OCT-1998.
XX	
PF	01-APR-1998; 98WO-DK000132.
XX	
PR	02-APR-1997; 97DK-00000376.
PR	18-APR-1997; 97US-0044624P.
PR	10-NOV-1997; 97DK-00001277.
PR	05-JAN-1998; 98US-0070488P.
XX	
PA	(STAT-) STATENS SERUM INST.
XX	
PI	Andersen P, Nielsen R, Rosenkrands I, Weldingh K, Rasmussen PB;
PI	Oettinger T, Florio W;
XX	
DR	WPI; 1998-542705/46.
DR	N-PSDB; AAV63939.
XX	
PT	New isolated mycobacteria polypeptides and nucleic acids - used for
PT	developing products for the diagnosis of or vaccination against
PT	mycobacterial infections, particularly tuberculosis.
XX	
PS	Claim 1; Page 200-202; 163pp; English.
XX	
CC	The present sequence represents a Mycobacterium tuberculosis protein.
CC	Products from the present invention, which describes protein fragments
CC	and nucleic acid fragments derived from M.tuberculosis, can be used in
CC	the detection of and prevention of mycobacterial infections. In
CC	particular, the proteins and nucleic acids can be used for the diagnosis
CC	of or vaccination against tuberculosis caused by M. tuberculosis, M.
CC	africanum or M. bovis

XX SQ Sequence 371 AA;

Query Match 100.0%; Score 1892; DB 2; Length 371;  
Best Local Similarity 100.0%; Pred. No. 2.2e-131;  
Matches 371; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MITMLWHAMPPELNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWT 60  
DB 1 MITMLWHAMPPELNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWT 60

QY 61 GGGSDKALAAATPMVWLQTASTQAKTRAMQATAQAAAAYTOAMATTPSLPEIAANHITQA 120  
DB 61 GGGSDKALAAATPMVWLQTASTQAKTRAMQATAQAAAAYTOAMATTPSLPEIAANHITQA 120

QY 121 VLTATNFFGINTPIALTMDYFIRMNQALAMEVYQAEATAVNTLFEKLEPMASILDPG 180  
DB 121 VLTATNFFGINTPIALTMDYFIRMNQALAMEVYQAEATAVNTLFEKLEPMASILDPG 180

QY 181 ASQSTTNPIFGMPSPGSSSTPVGQLPPAATQTLGOLGEMSGPMQQLTOPLOQVTSLSFQVG 240  
DB 181 ASQSTTNPIFGMPSPGSSSTPVGQLPPAATQTLGOLGEMSGPMQQLTOPLOQVTSLSFQVG 240

QY 241 GTGGGNPADEEAAQMGLLGTSPLSNHPPLAGGSGPSAGAGLLRAESLPGAGGSLTRTPLMS 300  
DB 241 GTGGGNPADEEAAQMGLLGTSPLSNHPPLAGGSGPSAGAGLLRAESLPGAGGSLTRTPLMS 300

QY 301 QLIKPVAPSVMPAAAAGSSATGGAAPVGAGAMGQAGGSTRPGLVAPAPLAQEREED 360  
DB 301 QLIKPVAPSVMPAAAAGSSATGGAAPVGAGAMGQAGGSTRPGLVAPAPLAQEREED 360

QY 361 DEDDWDDEDDW 371  
DB 361 DEDDWDDEDDW 371

RESULT 2  
AAY21946  
ID AAY21946 standard; protein; 371 AA.  
XX  
AC AAY21946;  
XX  
DT 06-SEP-1999 (first entry)  
XX  
DE Amino acid sequence of antigen RD1-ORF5.  
XX  
KW Immunogenic; Mycobacterium tuberculosis; immune response; infection;  
KW tuberculosis; fusion polypeptide; T-cell epitope; ESAT-6; MPT59; TB;  
KW pharmaceutical; vaccination; M. africanum; M. bovis; CFP7A; CFP30A;  
KW CFP7B; CFP19; CFP27; CFP30A; RD1-ORF; CFP10A; CFP16; CFP19; CFP23;  
KW CFP25A; CFP30B; CFP7B.  
XX  
OS Mycobacterium tuberculosis.  
XX  
PN W0924577-A1.  
XX  
PD 20-MAY-1999.  
XX  
PF 08-OCT-1998; 98WO-DK000438.  
XX  
PR 10-NOV-1997; 97DK-00001277.  
PR 05-JAN-1998; 98US-0070488P.  
PR 01-APR-1998; 98WO-DK000132.  
XX  
PA (STAT-) STATENS SERUM INST.  
XX  
PI Andersen P, Skjot R;  
XX  
XX WPI; 1999-347282/29.  
DR N-PSDB; AAX81046.  
XX  
XX New immunogenic fragment of Mycobacterium tuberculosis.  
PT  
XX

PS Example 2; Page 219-220; 265pp; English.

XX The invention describes a substantially pure immunogenic polypeptide  
CC fragment (I) from Mycobacterium tuberculosis that is able to evoke a  
CC protective immune response against infections by mycobacteria belonging  
CC to the tuberculosis complex. The invention provides a (1) fusion  
CC polypeptide comprising at least one polypeptide fragment (I) and at least  
CC one fusion partner; (2) a fusion polypeptide fragment comprising a T-cell  
CC epitope from M. tuberculosis protein ESAT-6, or MPT59 and a second  
CC different amino acid sequence from M. tuberculosis, and/or including a  
CC sequence which protects the first amino acid sequence from in vivo  
CC degradation or post-translational processing; (3) a nucleic acid fragment  
CC that encodes the above polypeptides. The polypeptides and nucleic acid  
CC are useful as pharmaceuticals, for diagnosis of and as antigens for  
CC vaccination against TB caused by Mycobacterium tuberculosis, africanum or  
CC bovis. The polypeptides are also useful for diagnosing ongoing or  
CC previous sensitization in an animal with bacteria belonging to the  
CC tuberculosis complex. The invention also describes the use of CFP7A or  
CC CFP30A or a T-cell epitope of for the induction of a strong immune  
CC response in a mammal; use of CFP7B, CFP19 or MPT59-ESAT6 or a T-cell  
CC epitope of for diagnosis of TB in a mammal by performing a DTH type skin  
CC test; use of CFP27, CFP30A, RD1-ORF2, RD1-ORF5, MPT59-ESAT6,  
CC ESAT6-MPT59, CFP10A, CFP19, CFP23, CFP25A, CFP30B, CFP7B or a T-  
CC cell epitope of for the preparation of an immunological composition; and  
CC for the preparation of a subunit vaccine

XX SQ Sequence 371 AA;

Query Match 100.0%; Score 1892; DB 2; Length 371;  
Best Local Similarity 100.0%; Pred. No. 2.2e-131;  
Matches 371; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MITMLWHAMPPELNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWT 60  
DB 1 MITMLWHAMPPELNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWT 60

QY 61 GGGSDKALAAATPMVWLQTASTQAKTRAMQATAQAAAAYTOAMATTPSLPEIAANHITQA 120  
DB 61 GGGSDKALAAATPMVWLQTASTQAKTRAMQATAQAAAAYTOAMATTPSLPEIAANHITQA 120

QY 121 VLTATNFFGINTPIALTMDYFIRMNQALAMEVYQAEATAVNTLFEKLEPMASILDPG 180  
DB 121 VLTATNFFGINTPIALTMDYFIRMNQALAMEVYQAEATAVNTLFEKLEPMASILDPG 180

QY 181 ASQSTTNPIFGMPSPGSSSTPVGQLPPAATQTLGOLGEMSGPMQQLTOPLOQVTSLSFQVG 240  
DB 181 ASQSTTNPIFGMPSPGSSSTPVGQLPPAATQTLGOLGEMSGPMQQLTOPLOQVTSLSFQVG 240

QY 241 GTGGGNPADEEAAQMGLLGTSPLSNHPPLAGGSGPSAGAGLLRAESLPGAGGSLTRTPLMS 300  
DB 241 GTGGGNPADEEAAQMGLLGTSPLSNHPPLAGGSGPSAGAGLLRAESLPGAGGSLTRTPLMS 300

QY 301 QLIKPVAPSVMPAAAAGSSATGGAAPVGAGAMGQAGGSTRPGLVAPAPLAQEREED 360  
DB 301 QLIKPVAPSVMPAAAAGSSATGGAAPVGAGAMGQAGGSTRPGLVAPAPLAQEREED 360

QY 361 DEDDWDDEDDW 371  
DB 361 DEDDWDDEDDW 371

RESULT 3  
AAB19844  
ID AAB19844 standard; protein; 368 AA.  
XX  
AC AAB19844;  
XX  
XX 05-MAR-2001 (first entry)  
DT  
XX Mycobacterium tuberculosis protein MTB3.  
XX  
XX MTB3; tuberculosis; BCG; vaccine; infection; diagnosis.  
XX

OS Mycobacterium tuberculosis.  
 XX WO200066157-A1.  
 XX 09-NOV-2000.  
 XX 04-MAY-2000; 2000WO-US012257.  
 XX 04-MAY-1999; 99US-0132505P.  
 XX (PUBL-) PUBLIC HEALTH RES INST NEW YORK.  
 XX PA Gennaro ML;  
 XX PI WPI; 2001-007153/01.  
 XX DR N-PSDB; AAA89037.  
 XX Novel polypeptide encoded by open reading frames present in Mycobacterium  
 PT tuberculosis genome and not by the BCG strain of M. bovis, useful as  
 PT vaccine and for diagnosing tuberculosis infection.  
 XX Claim 11; Fig 1; 35pp; English.  
 XX PS The present sequence is that of the Mycobacterium tuberculosis MTBN3  
 CC protein. This is 1 of 8 proteins, i.e. MTBN1-8 (see AAB19842-49), encoded  
 CC by 8 open reading frames (see AAA89035-42) identified as being present in  
 CC the genome of M. tuberculosis but absent from the genome of the BCG  
 CC strain of Mycobacterium bovis. MTBN1-8 represent reagents that are useful  
 CC in discriminating between M. tuberculosis and BCG and, in particular, for  
 CC diagnostic methods which discriminate between exposure of a subject to M.  
 CC tuberculosis and vaccination with BCG. The invention features these MTBN  
 CC polypeptides, functional fragments of them, DNA encoding them, vectors,  
 CC transformed cells, and diagnostic, therapeutic, and prophylactic  
 CC (vaccine) methods, including genetic vaccination methods  
 XX SQ Sequence 368 AA;  
 Query Match 99.3%; Score 1878; DB 4; Length 368;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-130;  
 Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 MLWHAMPPELNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWTGGG 63  
 Db 1 MLWHAMPPELNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWTGGG 60  
 QY 64 SDKALAAATPMVWVLTASTQAKTRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 123  
 Db 61 SDKALAAATPMVWVLTASTQAKTRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 120  
 QY 124 ATNFFGINTPIALTMDYFIRMWQAALAMEYQAEAVNTLFEKLEPMASILDPCASQ 183  
 Db 121 ATNFFGINTPIALTMDYFIRMWQAALAMEYQAEAVNTLFEKLEPMASILDPCASQ 180  
 QY 184 STTNPIFGMPSPGSSTPVQQLPPAATQTLGOLGEMSGPMQOOLTQPLQOQVTSLSFQVGGTG 243  
 Db 181 STTNPIFGMPSPGSSTPVQQLPPAATQTLGOLGEMSGPMQOOLTQPLQOQVTSLSFQVGGTG 240  
 QY 244 GGNPADEEAAQMGLLGTSPLSNHPLAGGSGPSAGAGLLRAESLPGAGGSLTRTPMLSQLI 303  
 Db 241 GGNPADEEAAQMGLLGTSPLSNHPLAGGSGPSAGAGLLRAESLPGAGGSLTRTPMLSQLI 300  
 QY 304 EKPVPSPVMPAAAGSATGGAAPVGAAGMGQAQSGGSTRPGLVAPAPLAQREDDDED 363  
 Db 301 EKPVPSPVMPAAAGSATGGAAPVGAAGMGQAQSGGSTRPGLVAPAPLAQREDDDED 360  
 QY 364 DWDEEDDW 371  
 Db 361 DWDEEDDW 368  
 RESULT 4  
 ADG15733  
 ID ADG15733 standard; protein; 368 AA.

XX AC ADG15733;  
 XX 26-FEB-2004 (first entry)  
 XX M tuberculosis Rv3873 protein for generating peptide epitopes.  
 XX antibacterial; antitubercular; tuberculostatic; vaccine;  
 KW diagnostic reagent; epitope; RD1 region; RD2 region;  
 KW Mycobacterium tuberculosis; Mycobacterium bovis; Mycobacterium africanum.  
 XX OS Mycobacterium tuberculosis.  
 XX WO2003093307-A2.  
 XX 13-NOV-2003.  
 XX 28-APR-2003; 2003WO-GB001815.  
 XX 27-APR-2002; 2002GB-00009723.  
 XX 27-APR-2002; 2002GB-00009724.  
 XX (UKEN-) UK SEC FOR ENVIRONMENT FOOD & RURAL AFF.  
 XX Cockle PJ, Vordermeier HM, Gordon SV, Hewinson RG;  
 XX WPI; 2003-903652/82.  
 XX New diagnostic reagents comprising a peptide having an epitope from  
 PT polypeptides Rv1986, Rv3878, Rv1983, Rv3873 or Rv3879 derived from RD1  
 PT and RD2 regions of Mycobacterium, useful as vaccines against  
 PT Mycobacterium infections.  
 XX Claim 19; SEQ ID NO 5; 77pp; English.  
 XX The invention relates to a novel diagnostic reagent comprising a peptide  
 CC epitope from a protein encoded by the RD1 or RD2 regions of the  
 CC Mycobacterium tuberculosis, M. bovis or M. africanum genome. The  
 CC diagnostic reagents comprised within the kit are selected so that they  
 CC are able to differentiate between M. bovis, M. tuberculosis or M.  
 CC africanum-infected mammals and mammals vaccinated against M. bovis, M.  
 CC tuberculosis or M. africanum. The polypeptide or its variant or fragment  
 CC is useful as a medicament provided that the polypeptide is not a MPT-64  
 CC polypeptide or a polypeptide encoded by the Rv1984C region of the M.  
 CC bovis, M. tuberculosis or M. africanum genomes. The nucleic acid which  
 CC encodes the polypeptide or the diagnostic reagents can be used as a  
 CC vaccine, especially against Mycobacterium infections. This sequence  
 CC corresponds to a protein used to derive the peptide epitopes used in the  
 CC method of the invention.  
 XX SQ Sequence 368 AA;  
 Query Match 99.3%; Score 1878; DB 7; Length 368;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-130;  
 Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 MLWHAMPPELNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWTGGG 63  
 Db 1 MLWHAMPPELNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWTGGG 60  
 QY 64 SDKALAAATPMVWVLTASTQAKTRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 123  
 Db 61 SDKALAAATPMVWVLTASTQAKTRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 120  
 QY 124 ATNFFGINTPIALTMDYFIRMWQAALAMEYQAEAVNTLFEKLEPMASILDPCASQ 183  
 Db 121 ATNFFGINTPIALTMDYFIRMWQAALAMEYQAEAVNTLFEKLEPMASILDPCASQ 180  
 QY 184 STTNPIFGMPSPGSSTPVQQLPPAATQTLGOLGEMSGPMQOOLTQPLQOQVTSLSFQVGGTG 243  
 Db 181 STTNPIFGMPSPGSSTPVQQLPPAATQTLGOLGEMSGPMQOOLTQPLQOQVTSLSFQVGGTG 240  
 QY 244 GGNPADEEAAQMGLLGTSPLSNHPLAGGSGPSAGAGLLRAESLPGAGGSLTRTPMLSQLI 303

```
Db 241 GGNPADEAAQWGLLGTSPLSNHPFLAGSGSPSAGAGLLRAESLPGAGSLTRTPLMSQLI 300
QY 304 EKPVPASVMPAAAAAGSSATGGGAAPVAGAGMGQGAQSGGSTREGLVAPAPLAQEREDEDED 363
Db 301 EKPVPASVMPAAAAAGSSATGGGAAPVAGAGMGQGAQSGGSTREGLVAPAPLAQEREDEDED 360
QY 364 DWDEDDW 371
Db 361 DWDEDDW 368

RESULT 5
AAW32384
ID AAW32384 standard; protein; 368 AA.
XX
AC AAW32384;
DT 13-JAN-1998 (first entry)
DE Mycobacterium tuberculosis antigen Tb37-FL.
KW Antigen; immunogen; vaccine; tuberculosis; non specific adjuvant;
KW skin testing; M.tuberculosis.
XX
OS Mycobacterium tuberculosis.
XX
FH Key Location/Qualifiers
FT Misc-difference 10 /note= "Any amino acid"
FT
FN WO9709429-A2.
XX
PD 13-MAR-1997.
XX
PF 30-AUG-1996; 96WO-US014675.
XX
PR 01-SEP-1995; 95US-00523435.
PR 22-SEP-1995; 95US-00532136.
PR 22-MAR-1996; 96US-00620280.
PR 05-JUN-1996; 96US-00658800.
PR 12-JUL-1996; 96US-00680573.
XX
PA (CORI-) CORIXA CORP.
XX
XX Read SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
PI Vedvick TH, Twardzik DR;
XX
XX WPI; 1997-192904/17.
XX
XX New immunogenic polypeptide(s) from soluble M. tuberculosis antigens -
FT useful for diagnosis of M. tuberculosis infection.
XX
XX Example 3; Page 159-161; 190pp; English.
XX
CC A new immunogenic polypeptide has been developed comprising an
CC immunogenic part of a soluble Mycobacterium tuberculosis antigen (or its
CC variant differing only in conservative substitutions and/or
CC modifications). The present sequence represents a M.tuberculosis antigen,
CC Tb37-FL. The immunogenic polypeptide can be used to diagnose
CC M.tuberculosis infection by forming complexes with specific antibodies in
CC the sample. Fragments of DNA encoding the immunogenic polypeptide can be
CC used as diagnostic primers or probes and agents that bind to the antigen,
CC especially monoclonal antibodies or equivalent polyclonal antibodies, are
CC also used for diagnosis
XX
SQ Sequence 368 AA;
Query Match 99.0%; Score 1873; DB 2; Length 368;
Best Local Similarity 99.7%; Pred. No. 5.5e-130;
Matches 367; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 MLWHAMPPELNTARLMAGAGPAPMLAAAAGQTLISAALDAQAVELTARLNSLGEAWTGGG 63
```

```
Db 1 MLWHAMPPELNTARLMAGAGPAPMLAAAAGQTLISAALDAQAVELTARLNSLGEAWTGGG 60
QY 64 SDKALAAATPMVWLQTAATAQKTRAMQATAQAAAAYTOAMATTPLSPLEIAANHITQAVLT 123
Db 61 SDKALAAATPMVWLQTAATAQKTRAMQATAQAAAAYTOAMATTPLSPLEIAANHITQAVLT 120
QY 124 ATNFFGINTIPALTENDYFIRMNQAALANEVYQAEAVNTLREKLEPMASILLDPGASQ 183
Db 121 ATNFFGINTIPALTENDYFIRMNQAALANEVYQAEAVNTLREKLEPMASILLDPGASQ 180
QY 184 STTNPIFGMPSPGSTPVGQLPPAATQTLGOLGEMSGPMQQLTOPLQOVTSLSFSQVGGTG 243
Db 181 STTNPIFGMPSPGSTPVGQLPPAATQTLGOLGEMSGPMQQLTOPLQOVTSLSFSQVGGTG 240
QY 244 GGNPADEAAQWGLLGTSPLSNHPFLAGSGSPSAGAGLLRAESLPGAGSLTRTPLMSQLI 303
Db 241 GGNPADEAAQWGLLGTSPLSNHPFLAGSGSPSAGAGLLRAESLPGAGSLTRTPLMSQLI 300
QY 304 EKPVPASVMPAAAAAGSSATGGGAAPVAGAGMGQGAQSGGSTREGLVAPAPLAQEREDEDED 363
Db 301 EKPVPASVMPAAAAAGSSATGGGAAPVAGAGMGQGAQSGGSTREGLVAPAPLAQEREDEDED 360
QY 364 DWDEDDW 371
Db 361 DWDEDDW 368

RESULT 6
AAW32452
ID AAW32452 standard; protein; 368 AA.
XX
AC AAW32452;
DT 09-JAN-1998 (first entry)
DE Mycobacterium tuberculosis antigen Tb37-FL.
XX
KW Antigen; immunogen; vaccine; tuberculosis; non specific adjuvant;
KW skin testing; M.tuberculosis.
XX
OS Mycobacterium tuberculosis.
XX
FH Key Location/Qualifiers
FT Misc-difference 10 /note= "Any amino acid"
FT
FN WO9709428-A2.
XX
PD 13-MAR-1997.
XX
PF 30-AUG-1996; 96WO-US014674.
XX
PR 01-SEP-1995; 95US-00523436.
PR 22-SEP-1995; 95US-00533634.
PR 22-MAR-1996; 96US-00620874.
PR 05-JUN-1996; 96US-00659683.
PR 12-JUL-1996; 96US-00680574.
XX
PA (CORI-) CORIXA CORP.
XX
XX Read SG, Skeiky YA, Dillon DC, Campos-Neto A, Houghton R;
PI Vedvick TH, Twardzik DR;
XX
XX WPI; 1997-192903/17.
XX
XX New immunogenic polypeptide(s) from Mycobacterium tuberculosis - are
FT useful in vaccines for prevention or treatment of tuberculosis, also for
FT diagnosis.
XX
XX Example 3; Page 146-147; 168pp; English.
XX
XX A new immunogenic polypeptide has been developed comprising an
```

CC immunogenic part of a soluble Mycobacterium tuberculosis antigen (or its  
 CC variant differing only in conservative substitutions and/or  
 CC modifications). The present sequence represents a M.tuberculosis antigen,  
 CC Tb37-FL The immunogenic protein, and fusion proteins containing one or  
 CC more of the proteins or one of the proteins plus ESAT-6, are useful in  
 CC vaccines, preferably when formulated with a non-specific adjuvant, to  
 CC induce an immune response against M.tuberculosis (for treatment or  
 CC prevention)  
 XX  
 XX Sequence 368 AA;

Query Match 99.0%; Score 1873; DB 2; Length 368;  
 Best Local Similarity 99.7%; Pred. No. 5.5e-130;  
 Matches 367; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 MLWHAMPPELNTARLMAGAGPAPMLAAAGWQTLNDAQAVELTARLNSLGEAWTGGG 63  
 Db 1 MLWHAMPPELNTARLMAGAGPAPMLAAAGWQTLNDAQAVELTARLNSLGEAWTGGG 60

QY 64 SDKALAAATPMVVLQSTQAKTRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 123  
 Db 61 SDKALAAATPMVVLQSTQAKTRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 120

QY 124 ATNFFGINTPIALTMDYFIRMNQALAMEVYQAEAVNTLFEKLEPMASILDPGASQ 183  
 Db 121 ATNFFGINTPIALTMDYFIRMNQALAMEVYQAEAVNTLFEKLEPMASILDPGASQ 180

QY 184 STTNPIFGMPSPGSSPPVQQLPPAATQTLGQLGEMSGPMQQLTOPLOQVTSLSFQVGGTG 243  
 Db 181 STTNPIFGMPSPGSSPPVQQLPPAATQTLGQLGEMSGPMQQLTOPLOQVTSLSFQVGGTG 240

QY 244 GGNPADEEAAQMGLLGTSPLSNHPLAGGSGPSAGAGLLRAESLPGAGGSLTRTPMSQLI 303  
 Db 241 GGNPADEEAAQMGLLGTSPLSNHPLAGGSGPSAGAGLLRAESLPGAGGSLTRTPMSQLI 300

QY 304 EKPVPASVMPAAAGSSATGGAAPVCGAGMGQAGQSGGSTRPGLVAPAPLAQEREDEDD 363  
 Db 301 EKPVPASVMPAAAGSSATGGAAPVCGAGMGQAGQSGGSTRPGLVAPAPLAQEREDEDD 360

QY 364 DWDEEDDW 371  
 Db 361 DWDEEDDW 368

RESULT 7  
 AAW64338  
 ID AAW64338 standard; protein; 368 AA.

AC AAW64338;  
 XX  
 XX 17-OCT-2003 (revised)  
 DT 09-NOV-1998 (first entry)  
 XX  
 XX Mycobacterium tuberculosis antigen Tb37-FL.  
 XX  
 XX Tuberculosis; infection; diagnosis; antigen; Tb37-FL.  
 XX  
 XX Mycobacterium tuberculosis; strain H37Rv.  
 XX  
 XX Key Location/Qualifiers  
 FT Misc-difference 10 /note= "unidentified"

PN WO9816645-A2.  
 XX  
 XX 23-APR-1998.  
 XX  
 XX 07-OCT-1997; 97WO-US018214.  
 XX  
 XX 11-OCT-1996; 96US-00729622.  
 PR 13-MAR-1997; 97US-00818111.  
 XX  
 XX (CORI-) CORIXA CORP.

XX Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;  
 PI Vedwick TS, Twardzik DR, Lodes MJ;  
 XX  
 XX WPI; 1998-251292/22.  
 DR  
 XX New isolated Mycobacterium tuberculosis polypeptides and DNA - used to  
 PT develop products for the detection of M. tuberculosis infection and  
 PT diagnosis of tuberculosis.  
 XX  
 XX Example 3; Page 143-144; 250pp; English.

XX This polypeptide comprises Mycobacterium tuberculosis antigen Tb37-FL. It  
 CC is encoded by genomic DNA isolated from a M. tuberculosis strain H37Rv  
 CC genomic library using a probe from clone Tb38-1 (see AAV44384). The  
 CC invention relates to compositions and methods for diagnosing  
 CC tuberculosis. It provides polypeptides (see AAW64291-W64379) comprising  
 CC an antigenic portion of a soluble M. tuberculosis antigen, or an  
 CC immunogenic portion of an M. tuberculosis antigen, as well as DNA  
 CC sequences encoding such polypeptides, recombinant expression vectors and  
 CC transformed or transfected host cells. Also claimed are methods and  
 CC diagnostic kits for detecting M. tuberculosis infection in a patient  
 CC using these polypeptides, antibodies or oligonucleotide probes and  
 CC primers, for the diagnosis of tuberculosis. (Updated on 17-OCT-2003 to  
 CC standardise OS field)  
 XX  
 XX Sequence 368 AA;

Query Match 99.0%; Score 1873; DB 2; Length 368;  
 Best Local Similarity 99.7%; Pred. No. 5.5e-130;  
 Matches 367; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 MLWHAMPPELNTARLMAGAGPAPMLAAAGWQTLNDAQAVELTARLNSLGEAWTGGG 63  
 Db 1 MLWHAMPPELNTARLMAGAGPAPMLAAAGWQTLNDAQAVELTARLNSLGEAWTGGG 60

QY 64 SDKALAAATPMVVLQSTQAKTRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 123  
 Db 61 SDKALAAATPMVVLQSTQAKTRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 120

QY 124 ATNFFGINTPIALTMDYFIRMNQALAMEVYQAEAVNTLFEKLEPMASILDPGASQ 183  
 Db 121 ATNFFGINTPIALTMDYFIRMNQALAMEVYQAEAVNTLFEKLEPMASILDPGASQ 180

QY 184 STTNPIFGMPSPGSSPPVQQLPPAATQTLGQLGEMSGPMQQLTOPLOQVTSLSFQVGGTG 243  
 Db 181 STTNPIFGMPSPGSSPPVQQLPPAATQTLGQLGEMSGPMQQLTOPLOQVTSLSFQVGGTG 240

QY 244 GGNPADEEAAQMGLLGTSPLSNHPLAGGSGPSAGAGLLRAESLPGAGGSLTRTPMSQLI 303  
 Db 241 GGNPADEEAAQMGLLGTSPLSNHPLAGGSGPSAGAGLLRAESLPGAGGSLTRTPMSQLI 300

QY 304 EKPVPASVMPAAAGSSATGGAAPVCGAGMGQAGQSGGSTRPGLVAPAPLAQEREDEDD 363  
 Db 301 EKPVPASVMPAAAGSSATGGAAPVCGAGMGQAGQSGGSTRPGLVAPAPLAQEREDEDD 360

QY 364 DWDEEDDW 371  
 Db 361 DWDEEDDW 368

RESULT 8  
 AAW81705  
 ID AAW81705 standard; protein; 368 AA.

XX AAW81705;  
 XX  
 XX 27-JAN-1999 (first entry)  
 DT  
 XX M. tuberculosis immunogenic polypeptide Tb37-FL.  
 DE  
 XX Tuberculosis; immunogenic; soluble; antigen; protective immunity; TB;  
 KW vaccine; pharmaceutical; infection; diagnosis.  
 KW

```
XX OS Mycobacterium tuberculosis.
XX FH Key Location/Qualifiers
XX FT Misc-difference 10
XX FT /label= unknown
XX FT
XX PN WO9816646-A2.
XX PD 23-APR-1998.
XX XX
XX PF 07-OCT-1997; 97WO-US018293.
XX XX
XX PR 11-OCT-1996; 96US-00730510.
XX PR 13-MAR-1997; 97US-00818112.
XX XX
XX PA (CORI-) CORIXA CORP.
XX XX
XX PI Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
XX PI Vedwick TS, Twardzik DR, Lodes MJ;
XX XX
XX DR WPI; 1998-261042/23.
XX XX
XX PT Immunogenic Mycobacterium tuberculosis polypeptide(s) and DNA - used to
XX FT develop products for the detection of M. tuberculosis infection and for
XX FT diagnosis, treatment and prevention of tuberculosis.
XX PS
XX PS Example 3B; Page 137-138; 230pp; English.
XX XX
XX CC This sequence represents an immunogenic portion of a soluble
XX CC Mycobacterium tuberculosis (MT) antigen which can be used in a method for
XX CC inducing protective immunity against tuberculosis (TB). This sequence can
XX CC be formulated into vaccines and/or pharmaceutical compositions for
XX CC immunising against M. tuberculosis infection or may be used for the
XX CC diagnosis of tuberculosis
XX XX
XX SQ Sequence 368 AA;
Query Match 99.0%; Score 1873; DB 2; Length 368;
Best Local Similarity 99.7%; Pred. No. 5.5e-130;
Matches 367; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 MLWHAMPPELNTARLMAGAGPAPMLAAAGWQTLSSAALDAQAVELTARLNSLGEAWTGGG 63
DB 1 MLWHAMPPEXNTARLMAGAGPAPMLAAAGWQTLSSAALDAQAVELTARLNSLGEAWTGGG 60
QY 64 SDKALAAATPMVWLQTAQKTRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 123
DB 61 SDKALAAATPMVWLQTAQKTRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 120
QY 124 ATNPFGINPTIALTEMDFYFIRMNQALAMEVYQAEAVNTLFEKLEPMASILDPGASQ 183
DB 121 ATNPFGINPTIALTEMDFYFIRMNQALAMEVYQAEAVNTLFEKLEPMASILDPGASQ 180
QY 184 STTNPIFGMPSGSGSTPVGQLPPAATQTLQGLGEMSGPMQQLTQPLQVTSLSFQVGGTG 243
DB 181 STTNPIFGMPSGSGSTPVGQLPPAATQTLQGLGEMSGPMQQLTQPLQVTSLSFQVGGTG 240
QY 244 GGNPADEAAQWGLLGTSPLSNHPHLAGGSGPSAGAGLLRAESLPAGAGSLTRTPMLMSQLI 303
DB 241 GGNPADEAAQWGLLGTSPLSNHPHLAGGSGPSAGAGLLRAESLPAGAGSLTRTPMLMSQLI 300
QY 304 EKPVAPSYMPPAAAAGSSATGGAAPVAGAMGGQAGSGGSTRPGLVAPAPLAQEREDEDD 363
DB 301 EKPVAPSYMPPAAAAGSSATGGAAPVAGAMGGQAGSGGSTRPGLVAPAPLAQEREDEDD 360
QY 364 DWDEEDDW 371
DB 361 DWDEEDDW 368
```

```
ID AAY38992 standard; protein; 368 AA.
XX AC AAY38992;
XX DT 05-NOV-1999 (first entry)
XX DE M. tuberculosis recombinant antigen protein Tb37-Fl.
XX XX
XX KW Antigen; diagnosis; detection; infection; antibody; immunisation;
XX KW vaccine; immunity.
XX OS Mycobacterium tuberculosis.
XX PN WO9942118-A2.
XX PD 26-AUG-1999.
XX PF 17-FEB-1999; 99WO-US003265.
XX XX
XX PR 18-FEB-1998; 98US-00024753.
XX PR 05-MAY-1998; 98US-00072596.
XX XX
XX PA (CORI-) CORIXA CORP.
XX XX
XX PI Reed SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;
XX PI Vedwick TS, Twardzik DR, Lodes MJ, Hendrickson RC;
XX XX
XX DR WPI; 1999-527416/44.
XX XX
XX PT New polypeptide comprising antigenic portions of M. tuberculosis.
XX PS Example 3; Page 177-179; 323pp; English.
XX CC This invention describes novel recombinant antigens and their encoding
XX CC nucleic acids derived from Mycobacterium tuberculosis. The novel
XX CC polypeptides are useful for detecting M. tuberculosis infection in a
XX CC biological sample by detecting antibodies which bind with the
XX CC polypeptides, and are useful as vaccines for immunizing against M.
XX CC tuberculosis infection. The new detection methods are needed as current
XX CC vaccination strategies do not provide 100% immunity
XX XX
XX SQ Sequence 368 AA;
Query Match 99.0%; Score 1873; DB 2; Length 368;
Best Local Similarity 99.7%; Pred. No. 5.5e-130;
Matches 367; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 MLWHAMPPELNTARLMAGAGPAPMLAAAGWQTLSSAALDAQAVELTARLNSLGEAWTGGG 63
DB 1 MLWHAMPPEXNTARLMAGAGPAPMLAAAGWQTLSSAALDAQAVELTARLNSLGEAWTGGG 60
QY 64 SDKALAAATPMVWLQTAQKTRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 123
DB 61 SDKALAAATPMVWLQTAQKTRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 120
QY 124 ATNPFGINPTIALTEMDFYFIRMNQALAMEVYQAEAVNTLFEKLEPMASILDPGASQ 183
DB 121 ATNPFGINPTIALTEMDFYFIRMNQALAMEVYQAEAVNTLFEKLEPMASILDPGASQ 180
QY 184 STTNPIFGMPSGSGSTPVGQLPPAATQTLQGLGEMSGPMQQLTQPLQVTSLSFQVGGTG 243
DB 181 STTNPIFGMPSGSGSTPVGQLPPAATQTLQGLGEMSGPMQQLTQPLQVTSLSFQVGGTG 240
QY 244 GGNPADEAAQWGLLGTSPLSNHPHLAGGSGPSAGAGLLRAESLPAGAGSLTRTPMLMSQLI 303
DB 241 GGNPADEAAQWGLLGTSPLSNHPHLAGGSGPSAGAGLLRAESLPAGAGSLTRTPMLMSQLI 300
QY 304 EKPVAPSYMPPAAAAGSSATGGAAPVAGAMGGQAGSGGSTRPGLVAPAPLAQEREDEDD 363
DB 301 EKPVAPSYMPPAAAAGSSATGGAAPVAGAMGGQAGSGGSTRPGLVAPAPLAQEREDEDD 360
QY 364 DWDEEDDW 371
DB 361 DWDEEDDW 368
```



Db 361 DWDEDDW 368

RESULT 10  
 AAY39135  
 ID AAY39135 standard; protein, 368 AA.  
 XX  
 AC AAY39135;  
 XX  
 XX 05-NOV-1999 (first entry)  
 XX  
 XX M. tuberculosis antigen TB37-FL amino acid sequence.  
 DE  
 DE  
 XX  
 XX Mycobacterium tuberculosis; M. tuberculosis; antigen; immunogen;  
 KW immunotherapy; diagnosis; immunisation; vaccine; infection;  
 KW immune response; skin test.  
 XX  
 XX Mycobacterium tuberculosis.  
 OS  
 XX WO9942076-A2.  
 PN  
 XX 26-AUG-1999.  
 XX  
 XX 17-FEB-1999; 99WO-US003268.  
 PF  
 XX 18-FEB-1998; 98US-00025197.  
 PR 05-MAY-1998; 98US-00072967.  
 XX  
 XX (CORI-) CORIXA CORP.  
 PA  
 XX Read SG, Skeiky YAW, Dillon DC, Campos-Neto A, Houghton R;  
 PI Vedvick TS, Twardzik DR, Lodes MJ, Hendrickson RC;  
 PT  
 XX WPI; 1999-527409/44.  
 DR  
 XX  
 XX New antigens from Mycobacterium tuberculosis useful in diagnostic skin  
 PT tests and protective or therapeutic vaccines or compositions.  
 XX  
 XX Example 3; Page 132-133; 299pp; English.  
 XX  
 XX The present invention describes polypeptides comprising an immunogenic  
 CC part of a Mycobacterium tuberculosis antigen (Ag). Also described are  
 CC vaccines and fusion protein containing M. tuberculosis Ag's. M.  
 CC tuberculosis Ag's, DNAs encoding them, derived fusion proteins and other  
 CC polypeptides fragments, can be used in pharmaceutical compositions or  
 CC vaccines to generate a protective or therapeutic immune response to M.  
 CC tuberculosis and as reagents in skin tests for diagnosis of tuberculosis.  
 CC Ag can induce proliferation of, or cytokine secretion by, T, B or natural  
 CC killer cells and/or macrophages in tuberculosis-immune subjects. AAZ19249  
 CC to AAZ19460 and AAY39083 to AAY39225 are used in the exemplification of  
 CC the present invention  
 XX  
 XX Sequence 368 AA;  
 SQ

Query Match	99.0%;	Score 1873;	DB 2;	Length 368;
Best Local Similarity	99.7%;	Pred. No. 5.5e-130;		
Matches 367;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;

Qy 4 MLHWAMPPELNTARLMAGAGPAPMLAAAGWQTLSAALDAQAVELTARLNSLGEAWTGGG 63

Db 1 MLHWAMPPEXNTARLMAGAGPAPMLAAAGWQTLSAALDAQAVELTARLNSLGEAWTGGG 60

Qy 64 SDKALAAATPMVWLQTASTQAKTRAWQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 123

Db 61 SDKALAAATPMVWLQTASTQAKTRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 120

Qy 124 ATNFFGINTPIALTENDYFIRMWQAALAMEVYQAEATAVNTLFEKLEPMASILDPGASQ 183

Db 121 ATNFFGINTPIALTENDYFIRMWQAALAMEVYQAEATAVNTLFEKLEPMASILDPGASQ 180

Qy 184 STTNPIFGMPSPGSTPPVQLPPAATQTILGQIGEMSGPMQQLTOPLQVTSLSFQVGGTG 243

Db 181 STTNPIFGMPSPGSTPPVQLPPAATQTILGQIGEMSGPMQQLTOPLQVTSLSFQVGGTG 240

39	IAAAKPMITWLOSA-----AEQTQQAHRQMASTPGMAVITENHITQAILATINF	91
128	FGINTPIALTENDYFIRMNQAALAMEYQAETAVENTLFEKLEPMASILDPCA-----	181
92	FGINWPIAFTEAGDFICWRTCTALAMNSYQAETILLNTAFQKLEPMAAILNPSSYPSPA	151
182	--SQ--STNTNIFGMPSPGSSTPVCQLPPAATQTGLGCGEMSGPMQOLTQPIQQVTSLS	237
152	LTSQVNQPTQMISGFSALPSTVLIQ-----QTVGVAAE-----LARPQQVKSKLEF	198
238	QVGCTGGGNPAD---EEAAQWGLLGTGPLSNHPLAGSGSPSAGAGLLRAESLPCAGGSL	293
199	SIDSTGVYTSRQGDTESAHRIGLFGASTLSHPLVGITGTITDTRELLCAESLPSAGSL	258
294	TRTPLMS--QLIEKPVPAS-----VMPAAAAGSSATGGAAPVCGAMGQGAQGGST	343
259	AWTPLMTFOFIIKSIAPPEPRQVMLP-----PWAASPGHNAQDGGTT	302

RESULT 12	
ADG15750	
ID	ADG15750 standard; protein; 100 AA.
XX	
AC	ADG15750;
XX	
DT	26-FEB-2004 (first entry)
XX	
DE	Rv3873 sequence fragment used to generate epitopes.
XX	
XX	antibacterial; antitubercular; tuberculostatic; vaccine;
KW	diagnostic reagent; epitope; RD1 region; RD2 region;
KW	Mycobacterium tuberculosis; Mycobacterium bovis; Mycobacterium africanum.
XX	
OS	Mycobacterium tuberculosis.
XX	
PN	WO2003093307-A2.
XX	
PD	13-NOV-2003.
XX	
XX	28-APR-2003; 2003WO-GB001815.
PF	
XX	
PR	27-APR-2002; 2002GB-00009723.
PR	27-APR-2002; 2002GB-00009724.
XX	
XX	(UKEN-) UK SEC FOR ENVIRONMENT FOOD & RURAL AFF.
PA	
XX	
PI	Cockle PJ, Vordermeier HM, Gordon SV, Hewinson RG;
XX	
DR	WPI; 2003-903652/82.
XX	
PT	New diagnostic reagents comprising a peptide having an epitope from
PT	polypeptides Rv1986, Rv3878, Rv1983, Rv3873 or Rv3879 derived from RD1
PT	and RD2 regions of Mycobacterium, useful as vaccines against
PT	Mycobacterium infections.
XX	
XX	
PS	Disclosure; SEQ ID NO 23; 77pp; English.
XX	
CC	The invention relates to a novel diagnostic reagent comprising a peptide
CC	epitope from a protein encoded by the RD1 or RD2 regions of the
CC	Mycobacterium tuberculosis, M. bovis or M. africanum genome. The
CC	diagnostic reagents comprised within the kit are selected so that they
CC	are able to differentiate between M. bovis, M. tuberculosis or M.
CC	africanum-infected mammals and mammals vaccinated against M. bovis, M.
CC	tuberculosis or M. africanum. The polypeptide or its variant or fragment
CC	is useful as a medicament provided that the polypeptide is not a MPI-64
CC	polypeptide or a polypeptide encoded by the Rv1984c region of the M.
CC	bovis, M. tuberculosis or M. africanum genomes. The nucleic acid which
CC	encodes the polypeptide or the diagnostic reagents can be used as a
CC	vaccine, especially against Mycobacterium infections. This sequence
CC	corresponds to a protein used to derive the peptide epitopes used in the
CC	method of the invention.
XX	
SQ	Sequence 100 AA;
XX	

Query Match	26.5%; Score 502; DB 7; Length 100;
Best Local Similarity	100.0%; Pred. No. 2.1e-29;
Matches 100; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	92 ATAAQAAYTQAMATTSLPEIAANHITQAVLTATNFFGINTIPALTENDYFIRMNQAA 151 
Db	1 ATQAQAAYTQAMATTSLPEIAANHITQAVLTATNFFGINTIPALTENDYFIRMNQAA 60 
Qy	152 LAMEVYQAEAVNTLFEKLEPMASILDPGASQSTTNPIFG 191 
Db	61 LAMEVYQAEAVNTLFEKLEPMASILDPGASQSTTNPIFG 100 
RESULT 13	
ABU34364	ID ABU34364 standard; protein; 445 AA.
XX	AC ABU34364;
XX	DT 19-JUN-2003 (first entry)
XX	DE Protein encoded by Prokaryotic essential gene #19891.
XX	KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX	OS Mycobacterium bovis.
XX	XX W0200277183-A2.
XX	XX 03-OCT-2002.
XX	XX 21-MAR-2002; 2002WO-US009107.
XX	XX 21-MAR-2001; 2001US-00815242.
XX	XX 06-SEP-2001; 2001US-00948993.
XX	XX 25-OCT-2001; 2001US-0342923P.
XX	XX 08-FEB-2002; 2002US-00072851.
XX	XX 06-MAR-2002; 2002US-0362699P.
XX	(ELIT-) ELITRA PHARM INC.
XX	PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX	PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Foreyth RA, Xu HH;
XX	XX WPI; 2003-0239926/02.
XX	DR N-PSDB; ACA38234.
XX	PT New antisense nucleic acids, useful for identifying proteins or screening
XX	PT for homologous nucleic acids required for cellular proliferation to
XX	PT isolate candidate molecules for rational drug discovery programs.
XX	XX Claim 25; SEQ ID NO 62288; 1766pp; English.
XX	XX The invention relates to an isolated nucleic acid comprising any one of
XX	XX the 6213 antisense sequences given in the specification where expression
XX	XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX	XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX	XX encoding a polypeptide whose expression is inhibited by the antisense
XX	XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX	XX polypeptide or its fragment whose expression is inhibited by the
XX	XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX	XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX	XX proliferation or the activity of a gene in an operon required for
XX	XX proliferation; (7) identifying a compound that influences the activity of
XX	XX the gene product or that has an activity against a biological pathway
XX	XX required for proliferation, or that inhibits cellular proliferation; (8)
XX	XX identifying a gene required for cellular proliferation or the biological
XX	XX pathway in which a proliferation-required gene or its gene product lies
XX	XX or a gene on which the test compound that inhibits proliferation of an
XX	XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX	XX compound's activity; (11) a culture comprising strains in which the gene
XX	XX product is overexpressed or underexpressed; (12) determining the extent





GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model.  
Run on: January 5, 2006, 20:12:50 ; Search time 164 Seconds  
(without alignment)  
1596.043 Million cell updates/sec

Title: US-09-872-505A-6  
Perfect score: 1892  
Sequence: 1 MITMLWHAMPPELNTARLWA.....PLAQREREDDDWDDEDDW 371

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 2166443 seqs, 705528306 residues  
Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot 05.80.\*  
1: uniprot\_prot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1887	99.7	371	Q7D4P4_MYCTU	Q7d4p4 mycobacteri
2	1878	99.3	368	Q7TVG3_MYCBO	Q7tv93 mycobacteri
3	1878	99.3	368	Q7TVG3_MYCBO	Q7tv93 mycobacteri
4	527	27.9	302	O33085_MYCLE	Q7txf6 mycobacteri
5	445	23.5	435	Q7TVX6_MYCBO	O53268 mycobacteri
6	440	23.3	435	Q7TVX6_MYCBO	O53268 mycobacteri
7	336.5	23.1	437	Q7TVX6_MYCBO	O53268 mycobacteri
8	32.5	22.9	518	Q73RJ3_MYCPA	Q73tj3 mycobacteri
9	30.5	22.8	434	Q7TVG3_MYCBO	F31500 mycobacteri
10	430.5	22.8	513	Q7TVG3_MYCBO	Q7u2c9 mycobacteri
11	430.5	22.8	513	Q7TVG3_MYCBO	Q7da35 mycobacteri
12	426.5	22.5	531	Q73U78_MYCPA	Q73u78 mycobacteri
13	423.5	22.4	504	Q73U78_MYCPA	Q73t11 mycobacteri
14	419	22.1	488	Q741E7_MYCPA	Q741e7 mycobacteri
15	408.5	21.6	550	Q73X20_MYCPA	Q73y20 mycobacteri
16	407.5	21.5	493	Q73X20_MYCPA	Q73td6 mycobacteri
17	404.5	21.4	518	Q7D9S7_MYCTU	Q7d9s7 mycobacteri
18	404.5	21.4	518	Q7D9S7_MYCTU	O6mx41 mycobacteri
19	404.5	21.4	518	Q7U1Y9_MYCBO	Q7u1y9 mycobacteri
20	400.5	21.2	696	Q741T4_MYCPA	Q741t4 mycobacteri
21	399.5	21.1	538	Q73TF3_MYCPA	Q73tf3 mycobacteri
22	398	21.0	585	Q73WR5_MYCPA	Q73wr5 mycobacteri
23	394.5	20.9	301	Q745B0_MYCPA	Q745b0 mycobacteri
24	383.5	20.3	393	Q79FK6_MYCTU	Q79fk6 mycobacteri
25	383.5	20.3	393	Q79FK6_MYCTU	Q79fk6 mycobacteri
26	382.5	20.2	393	Q7D7Y8_MYCTU	Q7d7y8 mycobacteri
27	382.5	20.2	399	Q7TZH8_MYCBO	Q7tzj3 mycobacteri
28	382.5	20.2	403	Q7D7X5_MYCTU	Q7tz88 mycobacteri
29	382	20.2	536	Q79FV8_MYCTU	Q7d7x5 mycobacteri
30	380	20.1	421	Q73ZU6_MYCPA	Q79fv8 mycobacteri
31	380	20.1	536	Q7U2D5_MYCBO	Q73zu6 mycobacteri
					Q7u2d5 mycobacteri

32	380	20.1	575	2	Q8VKNE_MYCTU	Q8vkn6 mycobacteri
33	379.5	20.1	399	2	Q79FJ7_MYCTU	Q79fj7 mycobacteri
34	376.5	19.9	463	1	PPE01_MYCTU	Q10892 mycobacteri
35	376.5	19.9	463	2	Q7U2V1_MYCBO	Q7u2v1 mycobacteri
36	376	19.9	350	2	Q7TZJ2_MYCBO	Q7tzj2 mycobacteri
37	376	19.9	423	2	Q7TZJ2_MYCBO	Q7tzj2 mycobacteri
38	376	19.9	423	2	Q7D7X9_MYCTU	Q7d7x9 mycobacteri
39	374	19.8	350	2	Q79FK5_MYCTU	Q79fk5 mycobacteri
40	374	19.8	363	2	Q7D7Y7_MYCTU	Q7d7y7 mycobacteri
41	373	19.7	473	2	Q79FH3_MYCTU	Q79fh3 mycobacteri
42	371	19.6	391	2	Q7D631_MYCTU	Q7d631 mycobacteri
43	369.5	19.5	385	2	Q79FLC_MYCTU	Q79fl6 mycobacteri
44	367	19.4	391	2	Q7TX76_MYCBO	Q7tx76 mycobacteri
45	366.5	19.4	385	2	Q7TZR7_MYCBO	Q7tzr7 mycobacteri

ALIGNMENTS

RESULT 1  
ID Q7D4P4\_MYCTU PRELIMINARY; PRT; 371 AA.  
AC Q7D4P4;  
DT 05-JUL-2004 (Tremblrel. 27, Created)  
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)  
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)  
DE PPE family protein.  
GN OrderedLocusNames=MT3987;  
OS Mycobacterium tuberculosis.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;  
OC Mycobacterium tuberculosis complex.  
OX NCBI\_TaxID=1773;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX STRAIN=CDC 1551 / Oshkosh;  
RX MEDLINE=22206494; PubMed=12218036;  
RX DOI=10.1128/JB.184.19.5479-5490.2002;  
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,  
RA Hickey B.K., Kolonay J.P., Nelson W.C., Umayam L.A., Ermolaeva M.D.,  
RA Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M.,  
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,  
RA Fraser C.M.;  
RA "Whole-genome comparison of Mycobacterium tuberculosis clinical and  
RT laboratory strains.";  
RL J. Bacteriol. 184:5479-5490 (2002).  
EMBL; AB000516; AAK48355.1; -; Genomic\_DNA.  
DR TIGR; MT3987; -;  
DR InterPro; IPR000030; Microbac\_PPE.  
DR Pfam; PF00823; PPE; 1;  
SQ SEQUENCE 371 AA; 37686 MW; CC25F5865131C79B CRC64;

Query Match 99.7%; Score 1887; DB 2; Length 371;  
Best Local Similarity 99.7%; Pred. No. 4.3e-98;  
Matches 370; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MITMLWHAMPPELNTARLWAGAGPAPMLAAAGWQTLSSAALDAQAVELTARLNSLGEAWT 60  
DB 1 MITMLWHAMPPELNTARLWAGAGPAPMLAAAGWQTLSSAALDAQAVELTARLNSLGEAWT 60  
QY 61 GGSGDKALAAATPMVWVLOTASTQAKTRAMQATAQAAAYTQAMATTPLSPLEIAANHITQA 120  
DB 61 GGSGDKALAAATPMVWVLOTASTQAKTRAMQATAQAAAYTQAMATTPLSPLEIAANHITQA 120  
QY 121 VLTFATNFGINTIPIALTENDYFIRWNOAALAMEVYQAEVTAQNTLFEKLEPMASILDPG 180  
DB 121 VLTFATNFGINTIPIALTENDYFIRWNOAALAMEVYQAEVTAQNTLFEKLEPMASILDPG 180  
QY 181 ASQSTTNPIFGMPSPGSSSTPVGQLPPAATQTLGCGEMSGPMQOLTOPLQOQVTSLSRQVG 240  
DB 181 ASQSTTNPIFGMPSPGSSSTPVGQLPPAATQTLGCGEMSGPMQOLTOPLQOQVTSLSRQVG 240

```
QY 241 GTGGNPADEAAQOMGLIGTSPLSNHPPLAGSGSPAGAGLILRAESILPGAGGSLTRTPILMS 300
DB 241 GTGGNPADEAAQOMGLIGTSPLSNHPPLAGSGSPAGAGLILRAESILPGAGGSLTRTPILMS 300
QY 301 OLIEKPVAPSVMPAAAGSSATGGGAAPVAGAGMGGAQGGGSTRPGLVAPAPLAQERRED 360
DB 301 OLIEKPVAPSVMPAAAGSSATGGGAAPVAGAGMGGAQGGGSTRPGLVAPAPLAQERRED 360
QY 361 DEDDDWDEDDW 371
DB 361 DEDDDWDEDDW 371

RESULT 2
QY Q7TVG3 MYCBO PRELIMINARY; PRT; 368 AA.
AC Q7TVG3
DT 01-OCT-2003 (TremBLrel. 25, Created)
DT 01-OCT-2003 (TremBLrel. 25, Last sequence update)
DE PPE FAMILY PROTEIN.
GN Name=PPE68; OrderedLocusNames=MD3903;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1765;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AP2122/97; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RX MEDLINE=22709107; PubMed=12788972; PubMed=9634230; DOI=10.1038/31159;
RA Garner T., Eiglimeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkins R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR EMBL; BX248347; CAD96089.1; -; Genomic_DNA.
DR InterPro; IPR000030; Microbac_PPE.
DR Pfam; PF00823; PPE; 1.
KW Complete proteome.
SQ SEQUENCE 368 AA; 37330 MW; D78F44095F658CA2 CRC64;

Query Match 99.3%; Score 1878; DB 2; Length 368;
Best Local Similarity 100.0%; Pred. No. 1.4e-97;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 MLWAMPPELNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWTGGG 63
DB 1 MLWAMPPELNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWTGGG 60
QY 64 SDKALAAATPMVWLQTAQTRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 123
DB 61 SDKALAAATPMVWLQTAQTRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 120
QY 124 ATNFFGINTPIALTMDYFIRMNQALAMEVYQAEAVNTLFEKLEPMASILDPGASQ 183
DB 121 ATNFFGINTPIALTMDYFIRMNQALAMEVYQAEAVNTLFEKLEPMASILDPGASQ 180
QY 184 STTNPIFGWPSGSGSTPVGQLPPAATQTLGQEMSGPMQQLTOPLOQVTSLSFQVGGTG 243
DB 181 STTNPIFGWPSGSGSTPVGQLPPAATQTLGQEMSGPMQQLTOPLOQVTSLSFQVGGTG 240
QY 244 GGNPADEAAQOMGLIGTSPLSNHPPLAGSGSPAGAGLILRAESILPGAGGSLTRTPILMSOLI 303
DB 241 GGNPADEAAQOMGLIGTSPLSNHPPLAGSGSPAGAGLILRAESILPGAGGSLTRTPILMSOLI 300
QY 304 EKPVAPSVMPAAAGSSATGGGAAPVAGAGMGGAQGGGSTRPGLVAPAPLAQERREDDED 363
DB 301 EKPVAPSVMPAAAGSSATGGGAAPVAGAGMGGAQGGGSTRPGLVAPAPLAQERREDDED 360
QY 364 DEDDDWDEDDW 371
DB 361 DEDDDWDEDDW 368

RESULT 4
```

```
DB 361 DEDDDWDEDDW 368

RESULT 3
QY Q79F92 MYCTU PRELIMINARY; PRT; 368 AA.
AC Q79F92
DT 05-JUL-2004 (TremBLrel. 27, Created)
DT 05-JUL-2004 (TremBLrel. 27, Last sequence update)
DE PPE FAMILY PROTEIN.
GN Name=PPE68; OrderedLocusNames=Rv3873;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1773;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eiglimeier K., Gas S., Barry C.E. III,
RA Tekaita F., Badcock K., Baeham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
RA Murphy J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulten J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
DR EMBL; BX842584; CAES5647.1; -; Genomic_DNA.
DR Tuberculist; RV3873; -.
DR InterPro; IPR000030; Microbac_PPE.
DR Pfam; PF00823; PPE; 1.
KW Complete proteome.
SQ SEQUENCE 368 AA; 37330 MW; D78F44095F658CA2 CRC64;

Query Match 99.3%; Score 1878; DB 2; Length 368;
Best Local Similarity 100.0%; Pred. No. 1.4e-97;
Matches 368; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 MLWAMPPELNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWTGGG 63
DB 1 MLWAMPPELNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWTGGG 60
QY 64 SDKALAAATPMVWLQTAQTRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 123
DB 61 SDKALAAATPMVWLQTAQTRAMQATAQAAAYTQAMATTPSLPEIAANHITQAVLT 120
QY 124 ATNFFGINTPIALTMDYFIRMNQALAMEVYQAEAVNTLFEKLEPMASILDPGASQ 183
DB 121 ATNFFGINTPIALTMDYFIRMNQALAMEVYQAEAVNTLFEKLEPMASILDPGASQ 180
QY 184 STTNPIFGWPSGSGSTPVGQLPPAATQTLGQEMSGPMQQLTOPLOQVTSLSFQVGGTG 243
DB 181 STTNPIFGWPSGSGSTPVGQLPPAATQTLGQEMSGPMQQLTOPLOQVTSLSFQVGGTG 240
QY 244 GGNPADEAAQOMGLIGTSPLSNHPPLAGSGSPAGAGLILRAESILPGAGGSLTRTPILMSOLI 303
DB 241 GGNPADEAAQOMGLIGTSPLSNHPPLAGSGSPAGAGLILRAESILPGAGGSLTRTPILMSOLI 300
QY 304 EKPVAPSVMPAAAGSSATGGGAAPVAGAGMGGAQGGGSTRPGLVAPAPLAQERREDDED 363
DB 301 EKPVAPSVMPAAAGSSATGGGAAPVAGAGMGGAQGGGSTRPGLVAPAPLAQERREDDED 360
QY 364 DEDDDWDEDDW 371
DB 361 DEDDDWDEDDW 368

RESULT 4
```

033085	MYCLE	PRELIMINARY;	PRT;	302	AA.
ID	033085;				
AC	033085;				
DT	01-JAN-1998	(T-EMBLrel. 05, Created)			
DT	01-JAN-1998	(T-EMBLrel. 29, Last sequence update)			
DT	01-FEB-2005	(T-EMBLrel. 29, Last annotation update)			
DE	Hypothetical protein MLC8628.14c (PPE-family protein).				
GN	Name=MLC8628.14c; OrderedAccession=ML0051;				
GC	Mycobacterium leprae.				
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;				
OC	Corynebacterineae; Mycobacteriaceae; Mycobacterium.				
OX	NCBI_TaxID=1769;				
ON	[1]				
ON	NUCLEOTIDE SEQUENCE.				
RP	MDLLINE=931188700; PubMed=8446027;				
RA	Eiglnieier K., Honore N., Woods S.A., Caudron B., Cole S.T.;				
RT	"Use of an ordered cosmid library to deduce the genomic organization				
RT	of Mycobacterium leprae.";				
RL	Mol. Microbiol. 7:197-206(1993).				
RN	[2]				
RN	NUCLEOTIDE SEQUENCE.				
RA	Eiglnieier K., Garnier T., De Rosai E., Fsihi H., Cole S.T.;				
RL	Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.				
RN	[3]				
RN	NUCLEOTIDE SEQUENCE.				
RC	STRAIN=TN;				
RC	MDLLINE=21128732; PubMed=11234002; DOI=10.1038/35059006;				
RA	Cole S.T., Eiglnieier K., Parkhill J., James K.D., Thomson N.R.,				
RA	Wheeler P.R., Honore N., Garnier T., Churcher C.M., Harris D.E.,				
RA	Mungall K.L., Basham D., Brown D., Chillingworth T., Connor R.,				
RA	Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,				
RA	Holroyd S., Hornby T., Jagels K., Maclean C., Maclean J., Moule S.,				
RA	Murphy L.D., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,				
RA	Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,				
RA	Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,				
RA	Barrell B.G.;				
RT	"Massive gene decay in the leprosy bacillus.";				
RL	Nature 409:1007-1011(2001).				
DR	EMBL; Y14967; CAA75201.1; -; Genomic DNA.				
DR	EMBL; AL583917; CAC29559.1; -; Genomic DNA.				
DR	PIR; T10033; T10033.				
DR	Leproma; ML0051; -.				
DR	InterPro; IPR000030; Microbac_PPE.				
DR	Pfam; PF00823; PPE; 1.				
KW	Complete proteome; Hypothetical protein.				
SEQ	SEQUENCE 302 AA; 32135 MW; EDAC4CC0AF3BF3D0 CRC64;				
Query Match 27.9%; Score 527; DB 2; Length 302;					
Best Local Similarity 45.1%; Pred. No. 6.1e-22;					
Matches 134; Conservative 31; Mismatches 78; Indels 54; Gaps 9;					
Qy	68	LAATPMVWLTQASTQAKTAMQATQAQAAYTQAMATTPSLPEIAANHITOAVLTATNF 127			
Db	39	IAAAKPMITWQSA-----AEQTTOAEAHQAMASTPGMAVITEMHIITQAILATINF 91			
Qy	128	FGINTIPALTMDYFIRMNQALAMVYQAEATVNTFLFKLPFMASILDPGA----- 181			
Db	92	FGINWAPIAFTAEAGDFICMTQTALAMNSYQAEITLNTAFQKLEPMAILNPSSVSPESA 151			
Qy	182	--SQ--STTNP1FGMPSPGSTPVGQLPPAATQTGLQGLGMSGPMQQLTQPLQVTSULFS 237			
Db	152	LTSQVNVQFTQMSGPSALPSTQVLQ-----QTVGVQAE-----LARPQQVKSLFT 198			
Qy	238	QVGGTGGGNPAD-----EBAQWGLLTGTSPLSNHPLAGSGPSAGAGLRAESLPCAGGSL 293			
Db	199	SIDSTGVVTSAQRGDTESAHRIGLFGASTLSHPLVIGITGTTTTRLICARSLFSASGSL 258			
Qy	294	TRTPLMS--QLIEKVPAS-----VMPAAAAGSSATGGAAPVGAGAMGQAGSGST 343			
Db	259	AWTPLMTQFLIDKSIAPRQVRMLP-----PWAAGSPGHNAQGGTT 302			
RESULT 5					





```
Db 303 PILGA-----LAAAVVPGVAG-----LAGVAGLAAPVAVCAAGAPAAALVGSVA 346
QY 327 PVGAGMGGAQSGSGSTRPGLVAPAP-----LAQER 357
Db 347 PVSGGVSPQARLVAPE-----APASTSVSLASDR 379

RESULT 8
Q73TJ3 MYCPA PRELIMINARY; PRT; 518 AA.
AC Q73TJ3;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=MAP3725;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB017240; AAS06275.1; -; Genomic_DNA.
DR InterPro; IPR00030; Microbac_PPE.
DR Pfam; PF00823; PPE; 1.
KW Complete proteome.
SQ SEQUENCE 518 AA; 52170 MW; BAC663B26F9CCF62 CRC64;

Query Match 22.8%; Score 432.5; DB 2; Length 518;
Best Local Similarity 32.6%; Pred. No. 2.2e-16;
Matches 143; Conservative 41; Mismatches 155; Indels 99; Gaps 15;

QY 1 MTTMLHMPPELNTARLMAGAPAPMLAAAGCQTLAALDAQAVELTARLNSL-GEAW 59
Db 1 MTAPIWMAFPEVHSAALLSGPGPGLASAGAWNSLAAYTSAAELTAVCEVQAGAW 60
QY 60 TGGGSKALAAATPMVWLTQASTQAKRAMQATAQAAAYTQAMATTPSLPEIAANHITQ 119
Db 61 QGSAESYVNAHAPYLAWLQASADSAGAAARHEAARAAATTAALAAAMPTLGEAANHATH 120
QY 120 AVLTATNFFGINTPIALTEMVDYFIRWMOALAMEVYQATVNTLFEKLEPMASILDLP 179
Db 121 AALQATNFFGINTPIAIVNEADY-ARMWVQAATTMTTYQ---AVSTAATAVATPQ-TMPAP 175
QY 180 GASQSTNPI-FCMPSFGSS-----TPVGQLPPAA-TQ-----TL 212
Db 176 QIAKSTAATIPFORPYPNPTNFSQVWADLWYDIPYSIETFGPVPDPANWTQLFQFNVTF 235
QY 213 GQLGEMSGPMQO-----LTQPL-----Q 230
Db 236 ANLAGTPAKLAQFNSPVSFLFNSWPTLLWVLDFTAGRIDLVTLKFLLEQLLYVVLGGM 295
QY 231 QVTSLSFQVGGTGGGNPADEAAQMGLLGTSPLSNHPLAGG-----SGPSAGALLRAE 284
Db 296 AVTSLGAAAGAAG-----LVGLAGLA--SPGLPTGAEMVPTAPPPTGATPAPTA 344
QY 285 SLPGAGSLTRTPILMSQLIEKVPVSPVMPA-AAAGSSATGGAAP-----VGAGMGQGA 337
Db 345 FLIGGPASAPATIPVSSAAAAAPTAPAPAPAAVAAAGSAAPPAGPGGPPYLVGGMRVSSAA 404
QY 338 QSGGSTRPGLVAPAPLAQ 355
Db 405 SASAQSRPKSAASAAAAA 422

RESULT 9
PPE46 MYCTU STANDARD; PRT; 434 AA.
ID PPE46_MYCTU
AC P31500; O53265;
```

```
DT 01-JUL-1993 (Rel. 26, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical PPE-family protein PPE46.
GN Name=ppe46; OrderedLocusNames=Rv3018c, MT3098/MT3101;
GN ORFNames=MIV012.32c;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1773;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III,
RA Tekla F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornaby T., Jagsis K., Krogh A., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "deciphering the biology of Mycobacterium tuberculosis from the
RL complete genome sequence.";
RL Nature 393:537-544 (1998).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RX DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.P., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL J. Bacteriol. 184:5479-5490 (2002).
RN [3]
RP NUCLEOTIDE SEQUENCE OF 160-374.
RC STRAIN=Isolate 50410;
RA Patki A.H., Dale J.W.;
RL Submitted (APR-1991) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the mycobacterial PPE family.
CC -!- CAUTION: In strain Oshkosh the gene for this protein is
CC interrupted in position 307 by an IS6110 element.
CC -!- CAUTION: Was originally (Ref.3) thought to be a dihydrofolate
CC reductase.
CC -!- CAUTION: Ref.3 sequence differs from that shown due to frameshifts
CC in positions 294, 337 and 355.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; BX842581; CAE55538.1; -; Genomic DNA.
CC EMBL; AE000516; AAK47427.1; ALT SEQ; Genomic DNA.
CC EMBL; AE000516; AAK47430.1; ALT SEQ; Genomic DNA.
CC EMBL; X59271; CAA41961.1; ALT_FRAME; Genomic DNA.
CC PIR; E70857; E70857.
CC TIGR; MT3098; -.
CC TIGR; MT3101; -.
CC TubercuList; Rv3018c; -.
CC InterPro; IPR00030; Microbac_PPE.
CC Pfam; PF00823; PPE; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 434 AA; 43029 MW; 41D673C4BD389DD6 CRC64;

Query Match 22.8%; Score 430.5; DB 1; Length 434;
```



Query Match 22.8%; Score 430.5; DB 2; Length 513;  
 Best Local Similarity 33.3%; Pred. No. 2.8e-16;  
 Matches 134; Conservative 44; Mismatches 155; Indels 69; Gaps 13;

QY 1 MITMLHAMPPELNTARLMAGAGPAPMLAAAGMOTLSAALDAQAVALTARLNSL-GEAW 59  
 DB 1 MAAPITWASPPVHSHALLSNGPGSLVAATAWQSLSAEYASTAAELSGLLGAVPGWAW 60

QY 60 TGGGSDKALAAATPMVWMLQTASTQAKTRAMQATAQAAYTQAMATTSLPBIANHITO 119  
 DB 61 QGSAEWYVAHLPYVWMLTQASADAAGAAACHEAAAAAYTTALAAMPTLAEAAAHVH 120

QY 120 AVLTATNFFGINTIPALTMDYFIRMWQALAMEVYQAEVAVNTLFEKLEPMASILD 178  
 DB 121 TVLVATNFFGINTIPALTMDYFIRMWQALAMEVYQAEVAVNTLFEKLEPMASILD 178

QY 179 PGASQSTTNPIFGMPSPGSSSTP-----VGQPPAATQTLQGLGEMSGPMQO----- 224  
 DB 180 FGGGAAT-----VGAVNPWQWLLALLQQLWNAVYTGFGWMLQLIWLQDPIGNSI 231

QY 225 -----LTPLOQVTS-----LFSQVGGTGGGNPADEEAAQMGLLGTSPLSNH 266  
 DB 232 KIILAFITNPIQALITYGELLFALGYQIFPNLVG-----PTWGMILSSPFL-L 279

QY 267 PLAGGSGPSAGAGL-----LRAESLPGAGGSLTRTPMLSQLIEKPVAPSVMPAAAGSSATG 323  
 DB 280 PAGLGLGLAAIAFLFIVLAPAVIPPAS-----TPLAAAV-----AAGSVMPAV---SMVAT 328

QY 324 GAAPVAGAGMGQAGGSGSTRPGLVAPAPLAQBEREDDDW 365  
 DB 329 GAGTAGAATPAAGAAPSAGAAPAPATASAYAVGGSGDW 370

RESULT 12  
 Q73U78 MYCPA  
 ID Q73U78 MYCPA PRELIMINARY; PRT; 531 AA.

AC Q73U78;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Hypothetical protein.  
 GN OrderedLocustNames=MAP3490;  
 OS Mycobacterium paratuberculosis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;  
 OC Mycobacterium avium complex (MAC).  
 OX NCBI\_TaxID=1770;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;  
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB017239; AAS06040.1; -; Genomic\_DNA.  
 DR InterPro; IPR000030; Microbac\_PPE.  
 DR Pfam; PF00823; PPE; 1.  
 KW Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 531 AA; 53742 MW; B392FE119582DC8C CRC64;

Query Match 22.5%; Score 426.5; DB 2; Length 531;  
 Best Local Similarity 34.9%; Pred. No. 4.9e-16;  
 Matches 137; Conservative 35; Mismatches 161; Indels 59; Gaps 15;

QY 1 MITMLHAMPPELNTARLMAGAGPAPMLAAAGMOTLSAALDAQAVALTARLNSL-GEAW 59  
 DB 7 MTAPIWASPPVHSHALLSNGPGSLVAATAWQSLSAEYASTAAELSGLLGAVPGWAW 66

QY 60 TGGGSDKALAAATPMVWMLQTASTQAKTRAMQATAQAAYTQAMATTSLPBIANHITO 119  
 DB 67 QGSAEWYVAHLPYVWMLTQASADAAGAAACHEAAAAAYTTALAAMPTLAEAAAHVH 126

QY 120 AVLTATNFFGINTIPALTMDYFIRMWQALAMEVYQAEVAVNTLFEKLEPMASILD 178

DB 127 GALVATNFFGINTIPALTMDYFIRMWQALAMEVYQAEVAVNTLFEKLEPMASILD 185

QY 179 PGASQSTTNPIFGMP-----SPGSSSTPVGQQLPAAATQTLG-----QLGEMSG-PMOQLTQPL 229

DB 186 STAAHDDHDEHGDHDDHDFDPLNQFVAQLRLFGIDWDVPVEGTLNGLPYEAYTSPA 245

QY 230 Q-----QVTSLSFSQVGGTGGGNPA-----DEEAAQMG-GLLGTSP-LS 264

DB 246 DPLMVVVRALBELFSDFOQFQALLOENPAAAFQITELVLLDWPHTLAQLASWLTPQQL 305

QY 265 NHPLAGGSGPSAGAGLLRAESLPGAGGSLTRTPMLSQLIEKPVAPSVMPAAAA---GSSAT 322

DB 306 LVPALVAAAPFG-----ALAGPAGVAGQPPLEPA-----PVAEPATPSAAAPGLPAT 352

QY 323 GGAAPVAGAGMGQAGGSGSTRPGLVAPAPLA 354

DB 353 AGATPIAASA-----AASGPAPAP---TPAPTA 377

RESULT 13  
 Q73T11 MYCPA  
 ID Q73T11 MYCPA PRELIMINARY; PRT; 504 AA.

AC Q73T11;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Hypothetical protein.  
 GN OrderedLocustNames=MAP3737;  
 OS Mycobacterium paratuberculosis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;  
 OC Mycobacterium avium complex (MAC).  
 OX NCBI\_TaxID=1770;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;  
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB017240; AAS06287.1; -; Genomic\_DNA.  
 DR InterPro; IPR000030; Microbac\_PPE.  
 DR Pfam; PF00823; PPE; 1.  
 KW Complete proteome; Hypothetical protein.  
 SQ SEQUENCE 504 AA; 49761 MW; 4DD3E220EB495932 CRC64;

Query Match 22.4%; Score 423.5; DB 2; Length 504;  
 Best Local Similarity 32.9%; Pred. No. 6.8e-16;  
 Matches 138; Conservative 39; Mismatches 140; Indels 103; Gaps 13;

QY 1 MITMLHAMPPELNTARLMAGAGPAPMLAAAGMOTLSAALDAQAVALTARLNSL-GEAW 59  
 DB 1 MTAPVWASPPVHSHALLSNGPGSLVAATAWQSLSAEYASTAAELSGLLGAVPGWAW 60

QY 60 TGGGSDKALAAATPMVWMLQTASTQAKTRAMQATAQAAYTQAMATTSLPBIANHITO 119  
 DB 61 EGPTEAYVAHAPYLAWLQSSADAAGAAACHEAAAAAYTTALAAMPTLPELATNATH 120

QY 120 AVLTATNFFGINTIPALTMDYFIRMWQALAMEVYQAEVAVNTLFEKLEPMASILD 177  
 DB 121 AALVATNFFGINTIPALTMDYFIRMWQALAMEVYQAEVAVNTLFEKLEPMASILD 176

QY 178 -----DPGAS-----QSTTNPIFGMPSPGSSSTPVGQQLPAAAT 209

DB 177 IVKSESSAGDTGDSGGGGGFGDLLSAWENFVWMLMDQLFGVDS-----PDLAS 227

QY 210 QTLGQGE-----MSGPMQQLTQPLQOVTSLFSQVGGTGGGNPADEEAAQMGLLGTSPLS 264

DB 228 AVAAFLANPSALFSAALLFALAYEIAFDLFFSPV-----ALLAAPFLP 271

QY 265 NHPLAGGSGPSAGAGLLRA-----BSLPGAGGSLTRTPMLSQLIEK 305

DB 272 FVLGAGLAGLAGLAGLAPPEPVPATVADVPATSPNQALPVAGLPSASTPAS----- 325

QY 306 FVAPSVMPAAAAAGSSATGGA-APVGAGAM-----GGAQSGGST--RPGLVAPAPLA 354

[illegible]

GenCore version 5.1.6  
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model  
Run on: January 5, 2006, 16:44:30 ; Search time 40 Seconds  
(without alignments)  
892.410 Million cell updates/sec

Title: US-09-872-505A-6  
Perfect score: 1892  
Sequence: 1 MITMLWHAMPPELNTARLMA.....PLAQRERDEDDWDEDDW 371

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 9621673 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 80:\*  
1: Pir1:\*  
2: Pir2:\*  
3: Pir3:\*  
4: Pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1878	99.3	368	2 G70802	probable PPE prote
2	527	27.9	302	2 T10033	hypothetical prote
3	430.5	22.8	434	2 E70857	probable PPE prote
4	430.5	22.8	513	2 D70836	probable PPE prote
5	404.5	21.4	518	2 F70831	probable PPE prote
6	383.5	20.3	393	2 G70929	probable PPE prote
7	382	20.2	536	2 F70835	probable PPE prote
8	379.5	20.1	403	2 H70931	probable PPE prote
9	376.5	19.9	463	2 H70750	probable PPE prote
10	376	19.9	423	2 B70931	probable PPE prote
11	374	19.8	350	2 H70929	probable PPE prote
12	373	19.7	473	2 F70813	probable PPE prote
13	371	19.6	391	2 D70922	probable PPE prote
14	369.5	19.5	385	2 H70503	probable PPE prote
15	365.5	19.3	391	2 A70663	probable PPE prote
16	365	19.3	556	2 D70940	probable PPE prote
17	362	19.1	408	2 G70925	probable PPE prote
18	360.5	19.1	413	2 F70560	probable PPE prote
19	356	18.8	572	2 T45392	hypothetical prote
20	352	18.6	391	2 B70625	probable PPE prote
21	351.5	18.6	365	2 E70929	probable PPE prote
22	348	18.4	463	2 C70931	probable PPE prote
23	344.5	18.2	394	2 A70504	probable PPE prote
24	343.5	18.2	380	2 A70846	probable PPE prote
25	331	17.5	468	2 B87158	PPE-family protein
26	326	17.2	423	2 C70582	probable PPE prote
27	325.5	17.2	438	2 S77655	hypothetical prote
28	323.5	17.1	468	2 B70932	probable PPE prote
29	322.5	17.0	406	2 E70675	probable PPE prote

ALIGNMENTS

RESULT 1

G70802  
probable PPE protein - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004  
C:Accession: G70802  
R:Colo, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skellon, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Suleston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; MUID:98295987; PMID:9634230  
A:Accession: G70802  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-368 <COL>  
A:Cross-references: UNIPROT:O69738; UNIPARC:UPI0000165388; GB:AL022120; GB:AL123456; NT1  
A:Experimental source: strain H37RV  
C:Genetics: PPE  
A:Gene: PPE

Query Match	99.3%	Score 1878;	DB 2;	Length 368;
Best Local Similarity	100.0%	Pred. No. 1.9e-104;		
Matches 368;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	4	MLWHAMPPELNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWTGGG	63	
Db	1	MLWHAMPPELNTARLMAGAGPAPMLAAAGWOTLSAALDAQAVELTARLNSLGEAWTGGG	60	
QY	64	SDKALAAATPMVWLOTASTQAKRAMQATAAAAYTOAWATTPSLPEIAANHITQAVLT	123	
Db	61	SDKALAAATPMVWLOTASTQAKRAMQATAAAAYTOAWATTPSLPEIAANHITQAVLT	120	
QY	124	ATNFFGINIPIALTENDYFIRWNOALAMEYVQAEAVNTLFEKLEPMASILDPCASQ	183	
Db	121	ATNFFGINIPIALTENDYFIRWNOALAMEYVQAEAVNTLFEKLEPMASILDPCASQ	180	
QY	184	STTNPIFGMPSPGSSSTPVQQLPPAATQTLGQLEMSGPMQQLTQPLQVTSLSFQVGGTG	243	
Db	181	STTNPIFGMPSPGSSSTPVQQLPPAATQTLGQLEMSGPMQQLTQPLQVTSLSFQVGGTG	240	
QY	244	GNPADERAAQMGLLTSTPLSNHPLAGGSPSAGALLRAESLPGAGGSLTRTPPLMSQLI	303	
Db	241	GNPADERAAQMGLLTSTPLSNHPLAGGSPSAGALLRAESLPGAGGSLTRTPPLMSQLI	300	
QY	304	EKEVAPSVMPAAAGSSATGGAAPVCGAGAMGQAGGSTRPGLVAPAPLAQERREDDDED	363	
Db	301	EKEVAPSVMPAAAGSSATGGAAPVCGAGAMGQAGGSTRPGLVAPAPLAQERREDDDED	360	
QY	364	DWDEDDW	371	



```
Db      276 FDRPETHMLQATWLAEENPOLLVAAITPAISGLGAVSGLAGLTGLVPPQPVPVVAPADPAV 335

Cy      309 PSVNAAAAGSSSATGAAPVGACAMGGCAQGSGSTRPLGVAP 350
        |:|:|:|:~::~||::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db      336 PTVLPLAGTAGTTPTTASPASAPAAGA-APGPAGTATATSASVP 376

RESULT 6
Gb       G70929 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: G70929
R:Cole, S.T.; Broesch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A:Authors: Sqares, R.; Sulatston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G. A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome A:Reference number: A70500; PMID:98295987; PMID:9634230
A:Accession: G70929
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-393 <COL>
A:A/Cross-references: UNIPROT:O53939; UNIPARC:UIP00001652AA; GB:AL022021; GB:AL123456; NI A:Experimental source: strain H37rv
C:Genetics:
A:Gene: PPE

Query Match          20.3%; Score 383.5; DB 2; Length 393;
Best Local Similarity 32.0%; Pred.No. 8.2e-16;
Matches 131; Conservative 38; Mismatches 156; Indels 85; Gaps 14;

Cy    8   AMPPELLNTARLMACGAGPAPMLAAAGHQTLSAALDAQVELTARINSL-GEAWTGCGSDK 66
D     5   ALPPEVNSVRMYFGPGSPMPVAASAANWGLAAELSSAATGYETVTQLSSSEGWLGPSAA 64
Qy    67   ALAAATPMVMVLQTPASTQAOKTRMQATAQAAYATQMAATTPLSPIEIAANHITQAVLATN 126
D     65   MAEA VAFYVAVMSRAAQAEQAQTQARA AAAAFAAAPTAATVPPLLIAANRASLMQLISTN 124
Qy    127  FFGINTIPIALTENDYFIRMNQNALANEVYTQAEATVNTLFKELEPMASILDP----- 179
D     125  VFQQNTSAIAAAEAQYG--EMWAQDSAAAMYAYAGSSASASAVTFSTPPQIANFTAQGTQA 183
Qy    180 -----GASOST-TNPITFGMFPSCSSSTPVGOLPPAATQTLLGQ---LGMSGSGPMQOLTQ--- 227
```

```

184 AAVATAAGTAQSTILTEMITGLPN-----ALQSLTSPQLQSSNGSPLSWLDWQLL 230
228 -----PQQVTSLSQVCG-----TGGCNPADEEAQAQGLLCTSPLSNHP 267
231 FGTPNFPTSISALLTDLQPYASFYNTGLPYFSIGMGNFTQSAKTLGLIGSAAPA 290
268 LAG-----GSGPAGAGLLRAESLPGAGGSLTRTPLMSQLIEKVAP----- 309
291 AAGDAAKGLPLGGMLGGGPVA-AGLGNAAV-----GKLSVPFVWSGGLPGSVTPGAAPL 345
310 -----SVMPAAAGSSATGGGAAPVCGAGAGCGAQSGLSTRPGLVAPAPLA 354
346 PVSTVSAAPAPG-SLIGGLPLAGAGGAGAPRYG--FRPTVMARPPFA 392

RESULT 7
F70835
probable PPE protein - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: F70835
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon,
; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

```

Db	181	TRLDVLSIGQLIRDLIDFIANPKYLFEPFEQGF-SPAVTIVVLVALVALQLYDFLWYPY	233
QV	203	-----QLPPAATQTTLGOLGEMSGPMQQLTQPLQQVTSFLFSQVGGTGGGNPDDEAAQMG	256



Db 240 YASYGLLLPFTPTLSALTALSALHLNLPAGLLPAAALG-----PGDQGANLA 293  
QY 257 LGTSPLSNHPAGSGAGLRAESLPAG--GSLTRTPMSQLIEKPVAFSVMPA 314  
Db 294 VAVTATAAVP--CGSPPTSNP-----APAAPSSNVGSSAAPGISYAVPGLAPGVSSG 347  
QY 315 AAGSSATGGAAPVAGAGMGQAGQSGSTRPGL 347  
Db 348 PRAGTKSPDTAADTLATA-----GAARPGL 372

RESULT 10  
B70931  
probable PPE protein - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004  
C:Accession: B70931  
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; MUID:98295987; PMID:9634230  
A:Accession: B70931  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-423 <COL>  
A:Cross-references: UNIPROT:O53950; UNIPARC:UPI00000D5E6B; GB:AL022021; GB:AL123456; NID:98295987  
A:Experimental source: strain H37RV  
C:Genetics:  
A:Gene: PPE

Query Match 19.9%; Score 376; DB 2; Length 423;  
Best Local Similarity 31.7%; Pred. No. 2.5e-15;  
Matches 128; Conservative 35; Mismatches 153; Indels 88; Gaps 11;  
QY 9 MPPELNTARLMAGAGPAPMLAAAGWQTLSSAALDAQAVELTARLNSLGEAWTGGGSKAL 68  
Db 6 LPPEINSGRMVTFGPGPMLAAATAWDGLAVELHATAAGYASBELSALTGAWSGFSSTMA 65  
QY 69 AAATPMVWLQTASTQAKTRAMQATAQAAYTQAMATTPSLPEIAANHITQAVLTATNPF 128  
Db 66 SAAAPYVAMNSATVAHAEIAGRAQALAIAYEAAFAATVPPIAANRAQLMVLIIATNIF 125  
QY 129 GINTIPALTENDYFIRMWNOALAMEVYQ----- 158  
Db 126 QONTPALIMTEAQY-MEMWAQDAANMYGAGSSATASRTAFTEPPQTTHHGQLGQSSA 184  
QY 159 -AETA-----VNTLFEKL-----EPMASILDPGASQSTT-----NPIFGMP 193  
Db 185 VAQTAATAAGNLQSAPFQLLSAVPRALQGLALPTASQSASATPQWVTDLGNLSLTLGGA 244  
QY 194 SPSSSTPVGQLPAAATQTLGQEMSGPMQOLTQPLQVTSLSQVGG---TCGGNPAD 250  
Db 245 VTGPYTPPGVLPSPGVPYL--LGTSQ---VLVTQNGQGVSAIIGKIGKIPKIGTALAPLAE 299  
QY 251 EAAQMGILGTPSLSNHPLAGSGPSAG---AGLLRAESLPAGGSLTRTPMSQLIEKPV 307  
Db 300 FALHTPLIGSEGL-----GGGSVAGIGRAGLVKLSVPGW-----TV 338  
QY 308 APSVMPAAAAAGSSATGGAAPVAGAGMGQAGQSGSTRPGLVAPA 351  
Db 339 AAPEIFSPAAALQATRLAAAPIAATDAGGALLGGMALSGLAGRA 382

RESULT 11  
H70929  
probable PPE protein - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004  
C:Accession: H70929

R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; MUID:98295987; PMID:9634230  
A:Accession: H70929  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-350 <COL>  
A:Cross-references: UNIPROT:O53940; UNIPARC:UPI00001652AB; GB:AL022021; GB:AL123456; NID:98295987  
A:Experimental source: strain H37RV  
C:Genetics:  
A:Gene: PPE

Query Match 19.8%; Score 374; DB 2; Length 350;  
Best Local Similarity 32.9%; Pred. No. 2.6e-15;  
Matches 124; Conservative 42; Mismatches 149; Indels 62; Gaps 13;  
QY 8 AMPELNTARLMAGAGPAPMLAAAGWQTLSSAALDAQAVELTARLNSL-GEAWTGGGSDK 66  
Db 5 ALPPEINSGRMVTCGPGSGPMLAAANADGVAVELGLATGYASVIAELTGAPVWGAASLS 64  
QY 67 ALAAATPMVWLQTASTQAKTRAMQATAQAAYTQAMATTPSLPEIAANHITQAVLTATN 126  
Db 65 MVAATPYVAMLSQAAARAEQAGMQAAAAAAYEAAFAVMTVPPVITANRVLVMTLIATN 124  
QY 127 FPGINTIPALTENDYFIRMWNOALAMEVYQAEATVNTLPEKLEPMAS-----I 176  
Db 125 FFGQNSAALAVAEAOY-AEMWAQDAVAMYGYAASASAS---RLIPFAAPPKTTNSAGVV 180  
QY 177 LDPGASQSTNP-----IFGMPSPGSTPTVGPQLPPAATQTLGQGENSGPMQQLTQPL 229  
Db 181 AQAVASVSWPNENDMWLVRLGSIPTERTIVRL-----LQGSYLATGMARFLTISA 233  
QY 230 QQVTSLSFGVGGTGGGNPADDEAAQMLLGTSPLSNHPLAGGSGPSAGAGLLRAESLPGA 289  
Db 234 QQLT--FGPGYTAGSGGAWYPTQPAGL-----GAGPAVSASLARAEVP--- 276  
QY 290 GGSLSLTPMSQLIEKPVAPSVMPAAAAAGSSATGGAAPVAGAGMGQ---AQSGGST--- 343  
Db 277 -GRLSVPP--SWAVAAP-AFAEKPEAGTPTMSVITGEASCGQGLLRGIPLARAGRRTGAF 332  
QY 344 -----RPGIVAPAPLA 354  
Db 333 AHRYGFRHSVITRSPSA 349

RESULT 12  
F70513  
probable PPE protein - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
C:Accession: F70513  
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; MUID:98295987; PMID:9634230  
A:Accession: F70513  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-473 <COL>  
A:Cross-references: UNIPARC:UPI00000C1594; GB:297559; GB:AL123456; NID:98295987; PMID:9634230  
A:Experimental source: strain H37RV  
C:Genetics:  
A:Gene: PPE

Query Match 19.7%; Score 373; DB 2; Length 473;  
Best Local Similarity 28.9%; Pred. No. 4.2e-15;



C;Accession: A70663  
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
 Nature 393, 537-544, 1998  
 A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A;Reference number: A70500; MUID:98295987; PMID:9634230  
 A;Accession: A70663  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-391 <COL>  
 A;Cross-references: UNIPROT:P95247; UNIPARC:UPI00000D5E93; GB:Z83860; GB:AL123456; NID:9  
 A;Experimental source: strain H37Rv  
 C;Genetics:  
 A;Gene: PPE

Query Match	19.3%	Score 365.5;	DB 2;	Length 391;
Best Local Similarity	30.3%	Pred. No. 9.5e-15;		
Matches 120;	Conservative 38;	Mismatches 165;	Indels 73;	Gaps 12;

  

QY	9	MPPELNTARLMAGAGPAPMLAAAAGWQTLSAALDQAQVELTARLNSLGEA-WTGGGSDKA	67
Db	8	LPPEINSARIYAGAGGGLPFMAAAWEGLAADLRASASSFDVIAAGLAAGPWSGSPASVAM	67
QY	68	LAATPMVWVLOTASTOAKTRAMQATAQAAAYTOAMATTPELPEIAANHITQAVLTATNF	127
Db	68	AGAAAAPYVGLSAAAGQAELSAQATAAATAFERAAATVHPAAVTANRVLGALVATNI	127
QY	128	FGINTIPIALTMDYPIRMNQAAALAMEVYQAEIVAVNTLFEKLEP-----	172
Db	128	LGONTPAIAATFEDY-VEWMAQDVGMVGYHAGAA--AVAETILTSPSPPLDLGLASQA	184
QY	173	-----MASILDPGASQSTTNPIFGMPS--PGSSTFVGQLPPAATQTQLGGE-----MS	219
Db	185	GAQLTGMATSVGAALSPIAEGAVEGVPAVVAAAQSVAAGLPVDAAALQVQAAAAYPASMLI	244
QY	220	GPMMQLTQPLQQTSLFSQVGGTGGGNPADEEAAQMGILLGTSPLSNHLGAGSGPSAGAG	279
Db	245	GPMMQLAQ-----MGTTANTAGLAGAAGAAAGAAADV-----PTFAGDIASGTGLGGAGG	293
QY	280	LLRAESLPGAGGS--LTRTPMSQLTEKVPAPSVMPAAAAAGSSATGGAA-----	326
Db	294	L-----GAGMSAEIIGKARLVGAMSVPTTWEGSVPARMASSAMAGLGMPAEVPAAGGP	346
QY	327	-----PVGAGAMGCGAQS-----GSTRPGLVAPAP	352
Db	347	MGMMPMPMGNGGAGAGMPAGMGRGGGANPHVVQARP	382

Search completed: January 5, 2006, 21:15:37  
 Job time : 43 secs

This Page Blank (uspro)